Available online at www.sciencedirect.com
ScienceDirect
Journal of Functional Analysis 254 (2008) 2470–2502
JOURNAL OF Functional Analysis
www.elsevier.com/locate/jfa

On the Small Ball Inequality in all dimensions

Dmitriy Bilyk, Michael T. Lacey *, Armen Vagharshakyan

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30030, USA

Received 10 September 2007; accepted 14 September 2007

Available online 21 February 2008

Communicated by J. Bourgain

Abstract

Let h_R denote an L^∞ normalized Haar function adapted to a dyadic rectangle $R \subset [0,1]^d$. We show that for choices of coefficients $\alpha(R)$, we have the following lower bound on the L^∞ norms of the sums of such functions, where the sum is over rectangles of a fixed volume:

$$n^{\frac{d-1}{2}-\eta}\Big\| \sum_{|R|=2^{-n}} \alpha(R) h_R(x) \Big\|_{L^\infty([0,1]^d)} \gtrsim 2^{-n} \sum_{|R|=2^{-n}} \big|\alpha(R)\big|, \quad \text{for some } 0<\eta<\frac{1}{2}.$$

The point of interest is the dependence upon the logarithm of the volume of the rectangles. With $n^{(d-1)/2}$ on the left above, the inequality is trivial, while it is conjectured that the inequality holds with $n^{(d-2)/2}$. This is known in the case of $d=2$ [Michel Talagrand, The small ball problem for the Brownian sheet, Ann. Probab. 22 (3) (1994) 1331–1354, MR 95k:60049], and a recent paper of two of the authors [Dmitriy Bilyk, Michael T. Lacey, On the Small Ball Inequality in three dimensions, Duke Math. J., (2006), in press, arXiv: math.CA/0609815] proves a partial result towards the conjecture in three dimensions. In this paper, we show that the argument of [Dmitriy Bilyk, Michael T. Lacey, On the Small Ball Inequality in three dimensions, Duke Math. J., (2006), in press, arXiv: math.CA/0609815] can be extended to arbitrary dimension. We also prove related results in the subjects of the irregularity of distribution, and approximation theory. The authors are unaware of any prior results on these questions in any dimension $d \geqslant 4$.

Keywords: Discrepancy function; Small ball inequality; Brownian Sheet; Haar functions; Littlewood–Paley Inequalities; Kolmogorov entropy; Mixed derivative

* Corresponding author.
E-mail addresses: bilyk@math.gatech.edu (D. Bilyk), lacey@math.gatech.edu (M.T. Lacey), armenv@math.gatech.edu (A. Vagharshakyan).

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2007.09.010

1. The Small Ball Conjectures

In this paper we will prove results in dimension four and higher in three separate areas, number theory, approximation theory, and probability theory: (a) the theory of irregularities of distribution, (b) the Kolmogorov entropy of spaces of functions with bounded mixed derivative, and (c) small deviation inequalities for the Brownian sheet. As far as the authors are aware, these are the first results on these questions which provide more information than that given by an average case analysis. Underlying these three results is a central inequality, the *Small Ball Inequality* for the Haar functions, which we state here. The related areas are addressed in the next section.

In one dimension, the class of dyadic intervals is $\mathcal{D} := \{[j2^k, (j+1)2^k)\colon\ j,k \in \mathbb{Z}\}$. Each dyadic interval has a left and right half, indicated below, which are also dyadic. Define the Haar functions

$$h_I := -\mathbf{1}_{I_{\text{left}}} + \mathbf{1}_{I_{\text{right}}}.$$

Note that this is an L^∞ normalization of these functions, which we will keep throughout this paper.

In dimension d, a *dyadic rectangle* is a product of dyadic intervals, thus an element of $\mathcal{D}^d$. We define a Haar function associated to R to be the product of the Haar functions associated with each side of R, namely

$$h_{R_1\times\cdots\times R_d}(x_1,\ldots,x_d) := \prod_{j=1}^{d} h_{R_j}(x_j).$$

This is the usual 'tensor' definition.

We will concentrate on rectangles with fixed volume and consider a local problem. This is the 'hyperbolic' assumption, that pervades the subject. Our concern is the following theorem and conjecture concerning a *lower bound* on the L^∞ norm of sums of hyperbolic Haar functions:

Small Ball Conjecture 1.1. *For dimension $d \geqslant 3$ we have the inequality*

$$2^{-n} \sum_{|R|=2^{-n}} \bigl|\alpha(R)\bigr| \lesssim n^{\frac{1}{2}(d-2)} \Bigl\| \sum_{|R|\geqslant 2^{-n}} \bigl|\alpha(R)h_R\bigr| \Bigr\|_\infty. \tag{1.2}$$

Average case analysis—that is passing through L^2—shows that we always have

$$2^{-n} \sum_{|R|=2^{-n}} \bigl|\alpha(R)\bigr| \lesssim n^{\frac{1}{2}(d-1)} \Bigl\| \sum_{|R|\geqslant 2^{-n}} \alpha(R)h_R \Bigr\|_\infty.$$

Namely, the constant on the right is bigger than in the conjecture by a factor of $\sqrt{n}$. We refer to this as the 'average case estimate,' and refer to improvements over this as a 'gain over the average case estimate.' Random choices of coefficients $\alpha(R)$ show that the Small Ball Conjecture is sharp.

In dimension $d = 2$, the conjecture was resolved by [14].[1]

[1] This result should be compared to [12], as well as [7,16].

Talagrand's Theorem 1.3. *For dimensions $d \geqslant 2$, we have*

$$2^{-n} \sum_{|R|=2^{-n}} |\alpha(R)| \lesssim \left\| \sum_{|R|=2^{-n}} \alpha(R) h_R \right\|_\infty. \tag{1.4}$$

Here, the sum on the right is taken over all rectangles with area at least 2^{-n}.

The main result of this note is the next theorem, which shows that there is a gain over the trivial bound in the Small Ball Conjecture in dimensions $d \geqslant 3$. In dimension $d = 3$, this result was proved in [3]. The three-dimensional result and its present extension build upon the method devised by [1]. As far as the authors are aware, this is the first 'gain over the average case bound' known in dimensions four and higher.

Theorem 1.5. *In dimension $d \geqslant 3$, there exists a number $\eta(d) > 0$ such that for all choices of coefficients $\alpha(R)$, we have the inequality*

$$n^{\frac{d-1}{2} - \eta(d)} \left\| \sum_{|R| \geqslant 2^{-n}} \alpha(R) h_R \right\|_\infty \gtrsim 2^{-n} \sum_{|R|=2^{-n}} |\alpha(R)|. \tag{1.6}$$

We take this theorem as basic to our study, and use its proof to derive results on the three other questions mentioned at the beginning of the introduction.

The principal difficulty in three and higher dimensions is that two dyadic rectangles of the same volume can share a common side length. Beck [1] found a specific estimate in this case, an estimate that is extended in [3]. In this note, the main technical device is the extension of this estimate, in the simplest instance, to arbitrary dimensions, see Lemma 5.2. This lemma, and its extension to longer products Theorem 8.3, is the main technical innovation of this paper. The value of η that we can get out of this line of reasoning appears to be of the order d^{-2}, imputing additional interest to the methods of proof used to improve this estimate. Indeed, many aspects of our analysis are suboptimal, and the most essential techniques necessary to optimize the arguments of this paper are yet to be discovered.

2. Related results

2.1. The L^∞ norm of the Discrepancy Function

In d dimensions, take $\mathcal{A}_N$ to be N points in the unit cube, and consider the Discrepancy Function

$$D_N(x) := \sharp \mathcal{A}_N \cap [\vec{0}, \vec{x}) - N |[\vec{0}, \vec{x})|. \tag{2.1}$$

Here, $[\vec{0}, \vec{x}) = \prod_{j=1}^d [0, x_j)$, that is a rectangle with antipodal corners being $\vec{0}$ and $\vec{x}$. Relevant norms of this function must tend to infinity, in dimensions 2 and higher. The canonical result of this type is the following estimate proved in [11].

K. Roth's Theorem 2.2. *We have the universal estimate*

$$\|D_N\|_2 \gtrsim (\log N)^{(d-1)/2},$$

with the implied constant depending only upon dimension.

For all $1 < p < \infty$, $\|D_N\|_p$ admits the same lower bound, a result in [13]. The endpoint estimates of $p = 1, \infty$ are however much harder, with definitive information known only in two dimensions. The method of proof of this theorem, and the L^p variants can be summarized as follows: Fix $2N \leqslant 2^n < N$, and just project the Discrepancy Function onto the (hyperbolic) Haar functions $\{h_R\colon |R| = 2^{-n}\}$. By the Bessel inequality, this provides a lower bound on the L^2 norm of D_N. This same method of proof, with the Littlewood–Paley inequalities replacing the Bessel inequality, can be used to prove the L^p lower bound, for $1 < p < \infty$. See [2].

At L^∞, guided by the sharpness of the Small Ball Conjecture, we pose the conjecture below, which represents a $\sqrt{\log N}$ gain over the lower bound proved by Roth.

The L^∞ norm of Discrepancy Function Conjecture 2.3. *In dimension $d \geqslant 3$, we have the lower estimate valid for all point sets $\mathcal{A}_N$,*

$$\|D_N\|_\infty \gtrsim (\log N)^{d/2}.$$

In dimension $d = 2$, this is the theorem of [12]. In dimension $d = 3$, [1,3] give partial information about this conjecture. In this paper, we can prove the following result, which appears to be new in dimensions $d \geqslant 4$.

Theorem 2.4. *In dimension $d \geqslant 3$ there is a positive $\eta = \eta(d) > 0$ for which we have the uniform estimate*

$$\|D_N\|_\infty \gtrsim (\log N)^{(d-1)/2+\eta}.$$

The proof of this result follows easily from the method of proof of Theorem 1.5, and will be presented below.

2.2. Metric entropy of mixed derivative Sobolev spaces

While the special structure of the Haar functions can be exploited to prove the Small Ball Conjecture, one would *not* anticipate that this special structure is in fact essential to the conjecture. Thus, we formulate a smooth variant of the Small Ball Conjecture.

Fix a continuous non-constant function φ, supported on $[-1/2, 1/2]$, and of mean zero. For a dyadic interval I, let

$$\varphi_I(x) = \varphi\left(\frac{x - c(I)}{|I|}\right)$$

be a translation and rescaling of φ so that it is supported on I. Then, for a dyadic rectangle $R = R_1 \times \cdots \times R_d$, set

$$\varphi_R(x_1, \ldots, x_d) = \prod_{j=1}^{d} \varphi_{R_j}(x_j).$$

Smooth Small Ball Conjecture 2.5. *For dimension $d \geqslant 3$ we have the inequality*

$$2^{-n} \sum_{|R|=2^{-n}} \left|\alpha(R)\right| \lesssim n^{\frac{1}{2}(d-2)} \left\| \sum_{|R|\geqslant 2^{-n}} \alpha(R)\varphi_R \right\|_\infty. \tag{2.6}$$

The implied constant depends upon dimension d and φ only.

In this direction, we will prove a result in the same spirit as our main theorem.

Theorem 2.7. *Suppose φ is continuous, supported on $[-1/2, 1/2]$, of mean zero, and such that $\langle \varphi, h_{[-1/2,1/2]} \rangle \neq 0$. For dimension $d \geqslant 3$, there is a positive $\eta = \eta(d)$ so that we have the inequality*

$$2^{-n} \sum_{|R|=2^{-n}} \left|\alpha(R)\right| \lesssim n^{\frac{1}{2}(d-1)-\eta} \left\| \sum_{|R|\geqslant 2^{-n}} \alpha(R)\varphi_R \right\|_\infty. \tag{2.8}$$

The implied constant depends upon φ.

With this theorem, we can establish new results on the metric entropy of certain Sobolev spaces of functions with mixed derivative in certain L^p spaces. In d dimensions, consider the map

$$\mathrm{Int}_d f(x_1, \dots, x_d) := \int_0^{x_1} \cdots \int_0^{x_d} f(y_1, \dots, y_d)\, dy_1 \cdots dy_d.$$

We consider this as a map from $L^p([0,1]^d)$ into $C([0,1]^d)$. Clearly, the image of Int_d consists of functions with L^p integrable mixed partial derivatives. Let us set

$$\mathrm{Ball}\big(MW^p\big([0,1]^d\big)\big) := \mathrm{Int}_d\big(\big\{f \in L^p\big([0,1]^d\big)\colon\ \|f\|_p \lesssim 1\big\}\big).$$

That is, this is the image of the unit ball of L^p. This is the unit ball of the space of functions with mixed derivative in L^p.

These sets are compact in $C([0,1]^d)$, and it is of relevance to quantify the compactness, through the device of *covering numbers*. For $0 < \epsilon < 1$, set $N(\epsilon, p, d)$ to be the least number N of points $x_1, \dots, x_N \in C([0,1]^d)$ so that

$$\mathrm{Ball}\big(MW^p\big([0,1]^d\big)\big) \subset \bigcup_{n=1}^{N} (x_n + \epsilon B_\infty).$$

Here, B_∞ is the unit ball of $C([0,1]^d)$. The task at hand is to uncover the correct order of growth of these numbers as $\epsilon \downarrow 0$. The case of $d = 2$ below follows from Talagrand [14], and the upper bound is known in full generality [5,16].

Conjecture 2.9. *For $d \geqslant 2$ one has the estimate*

$$\log N(\epsilon, p, d) \simeq \epsilon^{-1}(\log 1/\epsilon)^{d-1/2}, \quad \epsilon \downarrow 0.$$

It is well known [15] that results such as Theorem 2.7 can be used to give new lower bounds on these covering numbers.

Theorem 2.10. *For $1 \leqslant p < \infty$, and $d \geqslant 3$, there is an $\eta > 0$ for which we have*

$$\log N(\epsilon, p, d) \gtrsim \epsilon^{-1} (\log 1/\epsilon)^{d-1+\eta}.$$

We have concentrated on the case of one mixed derivative, but various results on fractional derivatives are also interesting. See for instance [4,9].

2.3. The Small Ball Inequality for the Brownian sheet

Perhaps, it is worthwhile to explain the nomenclature 'Small Ball' at this point. The name comes from the probability theory. Assume that $X_t : T \to \mathbb{R}$ is a canonical Gaussian process indexed by a set T. The *Small Ball Problem* is concerned with estimates of $\mathbb{P}(\sup_{t \in T} |X_t| < \varepsilon)$ as ε goes to zero, i.e. the probability that the random process takes values in an L^∞ ball of small radius. The reader is advised to consult a paper by Li and Shao [10] for a survey of this type of questions. A particular question of interest to us deals with the Brownian sheet, that is, a centered Gaussian process indexed by the points in the unit cube $[0,1]^d$ and characterized by the covariance relation $\mathbb{E} X_s \cdot X_t = \prod_{j=1}^d \min(s_j, t_j)$.

Kuelbs and Li [8] have discovered a tight connection between the Small Ball probabilities and the properties of the reproducing kernel Hilbert space corresponding to the process, which in the case of the Brownian sheet is $WM^2([0,1]^d)$, the space described in the previous subsection. Their result, applied to the setting of the Brownian sheet in [5], states that

Theorem 2.11. *In dimension $d \geqslant 2$, as $\varepsilon \downarrow 0$ we have*

$$-\log \mathbb{P}\big(\|B\|_{C([0,1]^d)} < \varepsilon\big) \simeq \varepsilon^{-2} (\log 1/\varepsilon)^{\beta} \quad \textit{iff} \quad \log N(\varepsilon) \simeq \varepsilon^{-1} (\log 1/\varepsilon)^{\beta/2}.$$

Thus, in agreement with Conjecture 2.9, the conjectured form of the aforementioned probability in this case is the following:

The Small Ball Conjecture for the Brownian sheet 2.12. *In dimensions $d \geqslant 2$, for the Brownian sheet B we have*

$$-\log \mathbb{P}\big(\|B\|_{C([0,1]^d)} < \varepsilon\big) \simeq \varepsilon^{-2} (\log 1/\varepsilon)^{2d-1}, \quad \varepsilon \downarrow 0.$$

In dimension $d = 2$, this conjecture has been resolved by Talagrand in the already cited paper [14], in which he actually proved Conjecture 2.5 for a specific function φ and used it to deduce the lower bound in the inequality above.[2] In higher dimensions, the upper bounds are established, see [5], and the previously known lower bounds miss the conjecture by a single power of the logarithm.

Theorem 2.10 can be translated into the following result on the Small Ball probability for the Brownian sheet:

[2] The work of Talagrand bears strong similarities to the prior work of [7,12]. The argument of Talagrand was subsequently clarified by [16] and [4].

Theorem 2.13. *In dimensions $d \geqslant 3$, there exists $\eta > 0$ such that for the Brownian sheet B we have*

$$-\log \mathbb{P}\big(\|B\|_{C([0,1]^d)} < \varepsilon\big) \gtrsim \varepsilon^{-2}(\log 1/\varepsilon)^{2d-2+\eta}, \quad \varepsilon \downarrow 0.$$

3. Notations and Littlewood–Paley inequality

Let $\vec{r} \in \mathbb{N}^d$ be a partition of n, thus $\vec{r} = (r_1, \dots, r_d)$, where the r_j are non-negative integers and $|\vec{r}| := \sum_t r_t = n$, which we refer to as the *length of the vector* $\vec{r}$. Denote all such vectors as $\mathbb{H}_n$. ('$\mathbb{H}$' for 'hyperbolic.') For vector $\vec{r}$ let $\mathcal{R}_{\vec{r}}$ be all dyadic rectangles R such that for each coordinate k, $|R_k| = 2^{-r_k}$.

Definition 3.1. We call a function f an r *function with parameter* $\vec{r}$ if

$$f = \sum_{R \in \mathcal{R}_{\vec{r}}} \varepsilon_R h_R, \quad \varepsilon_R \in \{\pm 1\}. \tag{3.2}$$

A fact used without further comment is that $f_{\vec{r}}^2 \equiv 1$.

As it has been already pointed out, the principal difficulty in three and higher dimensions is that the product of Haar functions is not necessarily a Haar function. On this point, we have the following

Proposition 3.3. *Suppose that $R_1, \dots, R_k$ are rectangles such that there is no choice of $1 \leqslant j < j' \leqslant k$ and no choice of coordinate $1 \leqslant t \leqslant d$ for which we have $R_{j,t} = R_{j',t}$. Then, for a choice of sign $\varepsilon \in \{\pm 1\}$ we have*

$$\prod_{j=1}^{k} h_R = \varepsilon h_S, \quad S = \bigcap_{j=1}^{k} R_k. \tag{3.4}$$

Proof. Expand the product as

$$\prod_{m=1}^{\ell} h_{R_m}(x_1, \dots, x_d) = \prod_{m=1}^{\ell} \prod_{t=1}^{d} h_{R_{m,t}}(x_t).$$

Here $\varepsilon_m \in \{\pm 1\}$. Our assumption is that for each t, there is exactly one choice of $1 \leqslant m_0 \leqslant \ell$ such that $R_{m_0,t} = S_t$. And moreover, since the minimum value of $|R_{m,t}|$ is obtained exactly once, for $m \neq m_0$, we have that $h_{R_{m,t}}$ is constant on S_t. Thus, in the t coordinate, the product is

$$h_{S_t}(x_t) \prod_{1 \leqslant m \neq m_0 \leqslant \ell} h_{R_{m,t}}(S_t).$$

This proves our lemma. $\square$

Remark 3.5. It is also a useful observation, that the product of Haar functions will have mean zero if the minimum value of $|R_{m,t}|$ is unique for at least one coordinate t.

Definition 3.6. For vectors $\vec{r}_j \in \mathbb{N}^d$, say that $\vec{r}_1, \dots, \vec{r}_J$ are *strongly distinct* iff for coordinates $1 \leqslant t \leqslant d$ the integers $\{r_{j,t}\colon\ 1 \leqslant j \leqslant J\}$ are distinct. The product of strongly distinct r functions is also an r function, which follows from 'the product rule' (3.3).

The r functions we are interested in are

$$f_{\vec{r}} := \sum_{R \in \mathcal{R}_{\vec{r}}} \operatorname{sgn}\bigl(\alpha(R)\bigr) h_R. \tag{3.7}$$

We recall some Littlewood–Paley inequalities, which are standard, and so we omit proofs.

Littlewood–Paley inequalities 3.8. *In one dimension, we have the inequalities*

$$\Bigl\| \sum_{I \subset \mathbb{R}} a_I h_I(\cdot) \Bigr\|_p \lesssim \sqrt{p} \Bigl\| \Bigl[\sum_{I \subset \mathbb{R}} a_I^2 \mathbf{1}_I(\cdot) \Bigr]^{1/2} \Bigr\|_p, \quad 2 < p < \infty. \tag{3.9}$$

Moreover, these inequalities continue to hold in the case where the coefficients a_I take values in a Hilbert space $\mathcal{H}$.

The growth of the constant is essential for us, in particular the factor $\sqrt{p}$ is, up to a constant, the best possible in this inequality. See [6,17]. That these inequalities hold for Hilbert space valued sums is imperative for applications to higher dimensional sums of Haar functions. The relevant inequality is as follows.

Theorem 3.10. *We have the inequalities below for hyperbolic sums of* r *functions in dimension* $d \geqslant 3$:

$$\Bigl\| \sum_{|\vec{r}|=n} f_{\vec{r}} \Bigr\|_p \lesssim (pn)^{(d-1)/2}, \quad 2 < p < \infty. \tag{3.11}$$

We recall a vector-valued harmonic analysis inequality.

Proposition 3.12. *Let $\mathcal{F}_j$ be a sigma field generated by dyadic rectangles in dimension* 2. *We then have*

$$\Bigl\| \Bigl[\sum_j \mathbb{E}(\varphi_j | \mathcal{F}_j)^2 \Bigr]^{1/2} \Bigr\|_p \lesssim p \Bigl\| \Bigl[\sum_j \varphi_j^2 \Bigr]^{1/2} \Bigr\|_p, \quad 2 < p < \infty. \tag{3.13}$$

Proof. This is one of many examples of a vector-valued inequality in the harmonic analysis literature. This particular inequality admits a simple proof by duality, recalled here for convenience.

Since $p > 2$, we can appeal to a duality argument. We can choose $g \in L^{(p/2)'}$ of norm one so that

$$\Bigl\| \sum_j \mathbb{E}(\varphi_j | \mathcal{F}_j)^2 \Bigr\|_{p/2} = \sum_j \bigl\langle \mathbb{E}(\varphi_j | \mathcal{F}_j)^2, g \bigr\rangle \leqslant \sum_j \bigl\langle \mathbb{E}\bigl(\varphi_j^2 \big| \mathcal{F}_j\bigr), g \bigr\rangle$$

$$= \sum_j \langle \varphi_j^2, \mathbb{E}(g|\mathcal{F}_j)\rangle \leqslant \sum_j \langle \varphi_j^2, \mathrm{M}g\rangle \leqslant \Big\| \sum_j \varphi_j^2 \Big\|_{p/2} \|\mathrm{M}g\|_{(p/2)'}$$

$$\lesssim \big((p/2)' - 1\big)^{-2} \Big\| \sum_j \varphi_j^2 \Big\|_{p/2}.$$

Here we have used Jensen's inequality and the self-duality of the conditional expectation operators. The operator $\mathrm{M}g$ is the (strong) maximal function on the plane, namely

$$\mathrm{M}g(x) = \sup_R \frac{\mathbf{1}_R}{|R|} \int_R |g(y)|\,dy,$$

where the supremum is over all dyadic rectangles R. This maps L^q into L^q for all $1 < q < \infty$, an inequality appealed to in the last line of the display above. Moreover, it is well known that the norm of the operator behaves as

$$\|\mathrm{M}\|_{q\to q} \lesssim (q-1)^{-2}, \quad 1 < q < 2. \qquad \square$$

4. Proof of Theorem 1.5

The proof of the theorem is by duality, namely we construct a function Ψ of L^1 norm about one, which is used to provide a lower bound on the L^∞ norm of the sum of Haar functions. The details of this argument are similar to those of [3].

The function Ψ will take the form of a Riesz product, but in order to construct it, we need some definitions. Fix $0 < \varepsilon < 1$ to be a small number, ultimately of order $1/d^2$. Define relevant parameters by

$$q = \lfloor a n^{\varepsilon} \rfloor, \qquad b = \frac{1}{4}, \tag{4.1}$$

$$\tilde{\rho} = a q^b n^{-(d-1)/2}, \qquad \rho = \sqrt{q}\, n^{-(d-1)/2}. \tag{4.2}$$

Here a is a small positive constant, we use the notation $b = \frac{1}{4}$ throughout, so as not to obscure those aspects of the argument that dictate this choice. $\tilde{\rho}$ is a 'false' L^2 normalization for the sums we consider, while the larger term ρ is the 'true' L^2 normalization. Our 'gain over the average case estimate' in the Small Ball Conjecture is $q^b \simeq n^{\varepsilon/4}$.

Divide the integers $\{1, 2, \ldots, n\}$ into q disjoint intervals of equal length $I_1, \ldots, I_q$, ordered from smallest to largest. Let $\mathbb{A}_t := \{\vec{r} \in \mathbb{H}_n\colon r_1 \in I_t\}$. Let

$$F_t := \sum_{\vec{r} \in \mathbb{A}_t} f_{\vec{r}}, \qquad H_n := \sum_{|R| = 2^{-n}} \alpha(R) h_R. \tag{4.3}$$

Here, the $f_{\vec{r}}$ are as in (3.7). The Riesz product is a 'short product':

$$\Psi := \prod_{t=1}^{q} (1 + \tilde{\rho} F_t).$$

One can view the $\tilde{\rho} F_t$ as a 'poor man's $\operatorname{sgn}(F_t)$,' in that the Riesz product above tends to weight the region where the functions F_t align. Note the subtle way in which the false L^2 normalization enters into the product. It means that the product is, with high probability, positive. And of course, for a positive function F, we have $\mathbb{E}F = \|F\|_1$, with expectations being typically easier to estimate. This heuristic is made precise below.

Proposition 3.3 suggests that we should decompose the product Ψ into

$$\Psi = 1 + \Psi^{\mathrm{sd}} + \Psi^{\neg}, \tag{4.4}$$

where the two pieces are the 'strongly distinct' and 'not strongly distinct' pieces. To be specific, for integers $1 \leqslant u \leqslant q$, let

$$\Psi_k^{\mathrm{sd}} := \tilde{\rho}^k \sum_{1 \leqslant v_1 < \cdots < v_k \leqslant q} \sum_{\vec{r}_t \in \mathbb{A}_{v_t}}{}^{\mathrm{sd}} \prod_{t=1}^{u} f_{\vec{r}_t},$$

where $\sum^{\mathrm{sd}}$ is taken to be over all $\vec{r}_t \in \mathbb{A}_{v_t}$, $1 \leqslant m \leqslant k$ such that

$$\text{the vectors } \{\vec{r}_t \colon\ 1 \leqslant m \leqslant k\} \text{ are strongly distinct.} \tag{4.5}$$

Then define

$$\Psi^{\mathrm{sd}} := \sum_{k=1}^{q} \Psi_k^{\mathrm{sd}}. \tag{4.6}$$

With this definition, it is clear that we have

$$\langle H_n, \Psi^{\mathrm{sd}} \rangle = \langle H_n, \Psi_1^{\mathrm{sd}} \rangle \gtrsim q^b \cdot n^{-\frac{d-1}{2}} \cdot 2^{-n} \sum_{|R| = 2^{-n}} |\alpha_R|, \tag{4.7}$$

so that q^b is our 'gain over the trivial estimate,' once we prove that $\|\Psi^{\mathrm{sd}}\|_1 \lesssim 1$ (estimate (4.14) below). Proving this inequality is the main goal of the technical estimates of the following lemma:

Lemma 4.8. *We have these estimates*:

$$\mathbb{P}(\Psi < 0) \lesssim \exp\bigl(-Aq^{1-2b}\bigr), \tag{4.9}$$

$$\|\Psi\|_2 \lesssim \exp\bigl(a' q^{2b}\bigr), \tag{4.10}$$

$$\mathbb{E}\Psi = 1, \tag{4.11}$$

$$\|\Psi\|_1 \lesssim 1, \tag{4.12}$$

$$\bigl\|\Psi^{\neg}\bigr\|_1 \lesssim 1, \tag{4.13}$$

$$\bigl\|\Psi^{\mathrm{sd}}\bigr\|_1 \lesssim 1. \tag{4.14}$$

Here, $0 < a' < 1$ *in* (4.10) *is a small constant, decreasing to zero as* a *in* (4.1) *goes to zero*; *and* $A > 1$ *in* (4.9) *is a large constant, tending to infinity as* a *in* (4.1) *goes to zero.*

Proof. We give the proof of the lemma, assuming our main inequalities proved in the subsequent sections.

Proof of (4.9). Using the distributional estimate (6.3) of Theorem 6.1 proved in Section 5, and the definition of Ψ we estimate

$$\begin{aligned}\mathbb{P}(\Psi < 0) &\leqslant \sum_{t=1}^{q} \mathbb{P}(\tilde{\rho} F_t < -1) \\ &= \sum_{t=1}^{q} \mathbb{P}\big(\rho F_t < -a^{-1} q^{1/2-b}\big) \\ &\lesssim \exp\big(-ca^{-2} q^{1-2b}\big).\end{aligned}$$

Proof of (4.10). The proof of this is detailed enough and uses the results of subsequent sections, so we postpone it to Section 6, Lemma 6.5 below.

Proof of (4.11). Expand the product in the definition of Ψ. The leading term is one. Every other term is a product

$$\prod_{k \in V} \tilde{\rho} F_k,$$

where V is a non-empty subset of $\{1, \ldots, q\}$. This product is in turn a linear combination of products of r functions. Among each such product, the maximum in the first coordinate is unique. This fact tells us that the expectation of these products of r functions is zero. So the expectation of the product above is zero. The proof is complete.

Proof of (4.12). We use the first two estimates of our lemma. Observe that

$$\begin{aligned}\|\Psi\|_1 &= \mathbb{E}\Psi - 2\mathbb{E}\Psi \mathbf{1}_{\Psi<0} \\ &\leqslant 1 + 2\mathbb{P}(\Psi < 0)^{1/2} \|\Psi\|_2 \\ &\lesssim 1 + \exp\big(-Aq^{1-2b}/2 + a' q^{2b}\big).\end{aligned}$$

We have taken $b = 1/4$ so that $1-2b = 2b$. For sufficiently small a in (4.1), we will have $A \gtrsim a'$. We see that (4.12) holds.

Indeed, Lemma 6.5 proves a uniform estimate, namely

$$\sup_{V \subset \{1,\ldots,q\}} \mathbb{E} \prod_{v \in V} (1 + \tilde{\rho} F_t)^2 \lesssim \exp\big(a' q^{2b}\big).$$

Hence, the argument above proves

$$\sup_{V \subset \{1,\ldots,q\}} \Big\| \prod_{t \in V} (1 + \tilde{\rho} F_t) \Big\|_1 \lesssim 1. \tag{4.15}$$

Proof of (4.13). The primary facts are (4.15) and Theorem 8.3; we use the notation devised for that theorem.

We use the triangle inequality, estimate (4.10) of Lemma 4.8, Hölder's inequality, with indices q^{2b} and $(q^{2b})' = q^{2b}/(q^{2b}-1)$, the inclusion–exclusion identity (8.2) and estimate (8.4) of Theorem 8.3 in the calculation below. Notice that we have

$$\sup_{V\subset\{1,\dots,q\}} \Big\| \prod_{t\in V}(1+\tilde\rho F_t)\Big\|_{(q^{2b})'} \leqslant \sup_{V\subset\{1,\dots,q\}} \Big\| \prod_{t\in V}(1+\tilde\rho F_t)\Big\|_1^{(q^{2b}-1)/q^{2b}} \times \Big\| \prod_{t\in V}(1+\tilde\rho F_t)\Big\|_2^{2q^{-2b}} \lesssim 1.$$

We now estimate

$$\begin{aligned}
\|\Psi^{\neg}\|_1 &\leqslant \sum_{G \text{ admissible}} \Big\| \tilde\rho^{|V(G)|}\,\mathrm{SumProd}\big(X(G)\big)\cdot \prod_{t\in\{1,\dots,q\}-V(G)}(1+\tilde\rho F_t)\Big\|_1\\
&\leqslant \sum_{G \text{ admissible}} \big\| \tilde\rho^{|V(G)|}\,\mathrm{SumProd}\big(X(G)\big)\big\|_{q^{2b}} \cdot \Big\| \prod_{t\in\{1,\dots,q\}-V(G)}(1+\tilde\rho F_t)\Big\|_{(q^{2b})'}\\
&\lesssim \sum_{G \text{ admissible}} \big\| \tilde\rho^{|V(G)|}\,\mathrm{SumProd}\big(X(G)\big)\big\|_{q^{2b}}\\
&= \sum_{v=2}^{q} \sum_{G:\ |V(G)|=v} \big\| \tilde\rho^{|V(G)|}\,\mathrm{SumProd}\big(X(G)\big)\big\|_{q^{2b}}\\
&\lesssim \sum_{v=2}^{q} \binom{q}{v} v^{2dv}\big[q^{C'}n^{-\eta}\big]^v\\
&\lesssim q^{C''}n^{-\eta} \lesssim n^{-\varepsilon'} \lesssim 1.
\end{aligned} \tag{4.16}$$

Proof of (4.14). This follows from (4.13) and (4.12), and the identity $\Psi = 1 + \Psi^{\mathrm{sd}} + \Psi^{\neg}$ together with the triangle inequality. □

5. The analysis of the coincidence

Following the language of J. Beck [1], a *coincidence* occurs if we have two vectors $\vec r \neq \vec s$ with e.g. $r_2 = s_2$. He observed that sums over products of r functions in which there are coincidences obey favorable L^2 estimates. We refer to (extensions of) this observation as the *Beck gain*. We introduce relevant notation for this situation. For $1 \leqslant k \leqslant d$ and $1 \leqslant t_1, t_2 \leqslant q$, set

$$\Phi_{t_1,t_2,k} := \sum_{\substack{\vec r\in\mathbb{A}_{t_1};\, \vec s\in\mathbb{A}_{t_2}\\ \vec r\neq\vec s\\ r_k=s_k}} f_{\vec r}\cdot f_{\vec s}. \tag{5.1}$$

Notice that due to our construction of the Riesz product, there are no coincidences in the first coordinate in the decomposition of Ψ, although the case $k = 1$ is important for the proof of the L^2 estimate (4.10). In the sum above, there are $2d-3$ free parameters among the vectors $\vec r$ and $\vec s$. That is, the pair of vectors $(\vec r, \vec s)$ are completely specified by their values in $2d-3$ coordinates. The following lemma suggests that these parameters behave as if they were orthogonal.

The simplest instance of the Beck gain 5.2. *We have the estimates below, valid for an absolute implied constant that is only a function of dimension* $d \geqslant 3$,

$$\sup \|\Phi_{t_1,t_2,k}\|_p \lesssim p^{d-1/2} \cdot n^{d-3/2}, \quad 2 \leqslant p < \infty, \tag{5.3}$$

where the supremum is taken over all $1 \leqslant k \leqslant d$ *and* $1 \leqslant t_1, t_2 \leqslant q$.

This estimate is smaller by $1/2$ power of n than what one might naively expect, and so we say that we have an average gain of $1/4$ power of n in the products above. (Here, the average is in reference to the two functions we form the product of.) This lemma, in dimension $d = 3$ appears in [3]. We will give an inductive proof of this estimate, that requires that we revisit the three-dimensional case. In the next section, we also derive other estimates from the one above.

The estimate above may admit an improvement, in that the power of p is perhaps too large by a single power, due to our use of Proposition 3.12. (There should also be a dependence upon q, but on this point, and in many others, the arguments of this paper are suboptimal, and so we do not pursue this point here.)

Conjecture 5.4. *We have the estimates below, valid for an absolute implied constant that is only a function of dimension* $d \geqslant 3$,

$$\sup \|\Phi_{t_1,t_2,k}\|_p \lesssim (pn)^{d-3/2}, \quad 2 \leqslant p < \infty. \tag{5.5}$$

Proof of Lemma 5.2. The proof is inductive on dimension. We shall suppress dependence on t_1, t_2. In fact, we shall prove the theorem for the quantity

$$\Phi_1 := \sum_{\substack{\vec{r} \neq \vec{s} \in \mathbb{H}_n \\ r_1 = s_1}} f_{\vec{r}} \cdot f_{\vec{s}}, \tag{5.6}$$

and the claimed statement will follow with only minor adjustments. To set up the induction, we need some definitions.

Definition 5.7. Given a set of r functions $\{f_{\vec{r}}\}$ and subset $\mathbb{C} \subset \mathbb{H}_{n_1} \times \cdots \times \mathbb{H}_{n_t}$, set

$$\operatorname{SumProd}(\mathbb{C}) = \sum_{(\vec{r}_1, \ldots, \vec{r}_t) \in \mathbb{C}} \prod_{s=1}^{t} f_{\vec{r}_s}.$$

Below, we will be interested in pairs and four-tuples of r functions. It is an important element of the argument, allowing us to run the induction, that we consider products of r functions where the vectors are in hyperbolic collections $\mathbb{H}_n$, for different values of n.

The main quantity we induct on is then

$$\mathcal{B}(d, n, p) = \sup_{\mathbb{B}} \big\| \operatorname{SumProd}(\mathbb{B}) \big\|_p, \quad d, n, p \geqslant 3. \tag{5.8}$$

Here, the supremum is formed over all $\mathbb{B} \subset \mathbb{H}_{n_1} \times \mathbb{H}_{n_2}$ and all r functions subject to these conditions:

- There is a coincidence in the first coordinate: For all $(\vec{r}, \vec{s}) \in \mathbb{B}$, we have $\vec{r} \neq \vec{s}$ and $r_1 = s_1$.
- $n_1, n_2 \leqslant n$. That is the lengths of the vectors $\vec{r}$ and $\vec{s}$ are permitted to be different.
- No other restriction is placed upon the pairs of vectors in $\mathbb{B}$.

Our main estimate on these quantities is as follows.

Lemma 5.9. *We have the inequality below valid for all dimensions $d \geqslant 3$,*

$$\mathcal{B}(d, n, p) \lesssim p^{d-1/2} n^{d-3/2}, \quad p, n \geqslant 3.$$

The inductive argument for Lemma 5.9 has the underlying strategy of reducing dimension by application of the Littlewood–Paley inequalities. But, this causes the collections of vectors to lose some of their symmetry. Regaining the symmetry causes us to introduce additional types of collections of vectors. Two of these collections are as follows:

$$\mathcal{C}(d, n, p) = \sup_{\mathbb{C}} \left\| \operatorname{SumProd}(\mathbb{C}) \right\|_p, \quad d, n, p \geqslant 3. \tag{5.10}$$

Here, the supremum is formed over all $\mathbb{C} \subset \mathbb{H}_{n_1} \times \mathbb{H}_{n_2}$ and all r functions subject to these conditions:

- There is a coincidence in the first coordinate: For all $(\vec{r}, \vec{s}) \in \mathbb{C}$, we have $\vec{r} \neq \vec{s}$ and $r_1 = s_1$.
- For all $(\vec{r}, \vec{s}) \in \mathbb{C}$, we have $r_2 > s_2$ and $r_3 < s_3$.
- $n_1, n_2 \leqslant n$.
- There is no other restriction on the pairs of vectors in $\mathbb{C}$.

The only difference between the present collections and the collections in $\mathcal{B}(d, n, p)$ is that in the present collections we assume locations of maximums in the second and third coordinates, thereby permitting application of the Littlewood–Paley inequalities in those two coordinates.

The second collection is less sophisticated. We simply assume that the maximum always occurs in say, the first coordinate. Define

$$\mathcal{D}(d, n, p) = \sup_{\mathbb{D}} \left\| \operatorname{SumProd}(\mathbb{D}) \right\|_p, \quad d, n, p \geqslant 3. \tag{5.11}$$

Here, the supremum is formed over all $\mathbb{D} \subset \mathbb{H}_{n_1} \times \mathbb{H}_{n_2}$ and all r functions subject to these conditions:

- There is a coincidence in the first coordinate: For all $(\vec{r}, \vec{s}) \in \mathbb{D}$, we have $\vec{r} \neq \vec{s}$ and $r_1 = s_1$.
- For all $(\vec{r}, \vec{s}) \in \mathbb{D}$, and all $2 \leqslant j \leqslant d$, we have $r_j \geqslant s_j$.
- $n_2 < n_1 \leqslant n$.

That is, we require that in each coordinate where there is a maximum, the maximum occurs in the vector $\vec{r}$.

Lemma 5.12. *We have the inequality below valid for all dimensions $d \geqslant 3$,*

$$\mathcal{C}(d, n, p), \mathcal{D}(d, n, p) \lesssim p^{d-1/2} \cdot n^{d-3/2}, \quad p, n \geqslant 3.$$

We turn to the proofs of Lemmas 5.9 and 5.12, and begin by explaining the logic of our induction. Let $\mathcal{B}(d)$ stand for the inequalities in Lemma 5.9 in dimension d, and likewise for $\mathcal{C}(d)$ and $\mathcal{D}(d)$. We prove:

- The inequalities $\mathcal{D}(d)$ for all dimensions d.
- The inequalities $\mathcal{B}(3)$ and $\mathcal{C}(3)$. At the same time, assuming $\mathcal{B}(d-1)$, $d \geqslant 4$, we prove $\mathcal{C}(d)$.
- Assuming $\mathcal{C}(d)$ and $\mathcal{D}(d)$, we prove $\mathcal{B}(d)$.

These clearly combine to prove the two lemmas, and so complete the proof of Lemma 5.2.

*The inequalities $\mathcal{D}(d)$.*The definition of $\mathcal{D}(d)$ permits the possibility of equality for a large number of coordinates of the two vectors. Let us exclude that case in this definition. Define

$$\mathcal{D}_{\neq}(d,n,p) = \sup_{\mathbb{D}} \big\| \mathrm{SumProd}(\mathbb{D}) \big\|_p, \quad d,n,p \geqslant 3, \tag{5.13}$$

where $\mathbb{D}$ is as in (5.11), but with the additional condition that for $2 \leqslant j \leqslant d$ we have $r_j > s_j$. Then, we are free to apply the Littlewood–Paley inequality in each of the coordinates from 2 to d.

Fix a collection of vectors $\mathbb{D}$ and a collection of r functions which achieves the supremum in (5.13). For this collection, and a choice of vector $\vec{\rho} \in \mathbb{N}^{d-1}$, let

$$\mathbb{D}_{\vec{\rho}} = \big\{ (\vec{r}, \vec{s}) \in \mathbb{D}\colon r_{j+1} = \rho_j,\ 1 \leqslant j \leqslant d-1 \big\}.$$

Of course there are at most $\lesssim n^{d-1}$ values of $\vec{\rho}$ for which the collection above is non-empty. Then,

$$\begin{aligned} \mathcal{D}_{\neq}(d,n,p) &\lesssim p^{(d-1)/2} \bigg\| \bigg[\sum_{\vec{\rho}} \mathrm{SumProd}(\mathbb{D}_{\vec{\rho}})^2 \bigg]^{1/2} \bigg\|_p \\ &\lesssim p^{(d-1)/2} n^{(d-1)/2} \sup_{\vec{\rho}} \bigg\| \sum_{\vec{\rho}} \mathrm{SumProd}(\mathbb{D}_{\vec{\rho}}) \bigg\|_p. \end{aligned}$$

But, the coordinate r_1 is completely specified in $\mathbb{D}_{\vec{\rho}}$, and therefore does not contribute to the last norm. And so the first coordinate of $\vec{s}$ is specified. Therefore, there are at most $d-2$ free choices of parameters in the vector s. By application of the Littlewood–Paley inequalities, we have

$$\mathcal{D}_{\neq}(d,n,p) \lesssim (pn)^{d-3/2}.$$

This is better than the claimed inequality.

If there is a set $J \subset \{2,\dots,d\}$ of coordinates for which $r_j = s_j$ for all $j \in J$, then after arbitrarily specifying these values, we have will be in position to apply the inequality $\mathcal{D}_{\neq}(d-|J|,n,p)$. This will clearly give a smaller estimate. As the number of possible choices for J is only a function of dimension, this completes the proof.

The bounds $\mathcal{B}(3)$ and $\mathcal{C}(3)$. Assuming $\mathcal{B}(d-1)$, $d \geqslant 4$, we prove $\mathcal{C}(d)$. In this section, we will prove the estimates for $\mathcal{C}(3)$. As well, we present the inductive proof of $\mathcal{C}(d)$ assuming $\mathcal{B}(d-1)$, for $d \geqslant 4$.

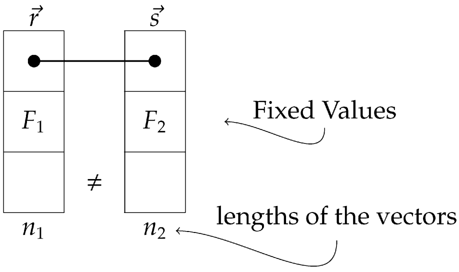

Fig. 1. The collections $\mathbb{M}$, with a coincidence in the top row, the second row taking fixed values, and no coincidence in the bottom row.

For the proof of $\mathcal{C}(3)$ there is an ancillary collection that we will have recourse to. Let

$$\mathcal{M}(n,p) = \sup_{\mathbb{M}} \left\| \operatorname{SumProd}(\mathbb{M}) \right\|_p, \tag{5.14}$$

where the supremum is formed over all choices of $\mathbb{M} \subset \mathbb{H}_{n_1} \times \mathbb{H}_{n_2}$ and all r functions subject to these conditions:

- $\vec{r}, \vec{s}$ are three-dimensional vectors.
- There is a coincidence in the first coordinate: For all $(\vec{r}, \vec{s}) \in \mathbb{C}$, we have $\vec{r} \neq \vec{s}$ and $r_1 = s_1$.
- The second coordinates are fixed: There are integers F_1, F_2 so that for all $(\vec{r}, \vec{s}) \in \mathbb{M}$ we have $r_2 = F_1$ and $s_2 = F_2$.
- There is no coincidence in the third coordinate: For all $(\vec{r}, \vec{s}) \in \mathbb{M}$ we have $r_3 \neq s_3$.
- $n_1, n_2 \leqslant n$.

See Fig. 1 for an illustration of this collection.We remark that in the case $n_1 \neq n_2$, a coincidence can occur in the third coordinate, a case that will come up below.

Lemma 5.15. *We have the inequalities*

$$\mathcal{M}(n,p) \lesssim \sqrt{p} \cdot \sqrt{n}. \tag{5.16}$$

Proof. Notice that the value of the maximum in the third coordinate completely specifies the pair of vectors $(\vec{r}, \vec{s})$. Therefore, one application of the Littlewood–Paley inequalities completes the proof. For any collection $\mathbb{M}$ as above, let $\mathbb{M}_a$ be the $(\vec{r}, \vec{s}) \in \mathbb{M}$ where the maximum in the third coordinate is a, $\max\{r_3, s_3\} = a$. Note that this can only consist, at most, of two pairs of vectors,

$$\left\| \operatorname{SumProd}(\mathbb{M}) \right\|_p \lesssim \sqrt{p} \left\| \sum_a \operatorname{SumProd}(\mathbb{M}_a)^2 \right\|_p /2^{1/2} \lesssim \sqrt{p} \cdot \sqrt{n}. \qquad \square$$

Fix a dimension $d \geqslant 3$. Let $\mathbb{B}$ be the collection which satisfies the conditions associated with (5.8) that contains $\mathbb{C}$. We introduce a conditional expectation into the argument, to gain some additional symmetry. Let $\mathcal{F}_{a,b}$ be the dyadic sigma field in the second and third coordinates generated by dyadic rectangles of side lengths 2^{-a-1} and 2^{-b-1}, respectively.

We have this equality:

$$\sum_{\substack{(\vec{r},\vec{s})\in\mathbb{C}\\ r_2=a,\, s_3=b}} f_{\vec{r}}\cdot f_{\vec{s}} = \mathbb{E}\Bigg(\sum_{\substack{(\vec{r},\vec{s})\in\mathbb{B}\\ r_2=a,\, s_3=b}} f_{\vec{r}}\cdot f_{\vec{s}} \big| \mathcal{F}_{a,b}\Bigg) - \operatorname{SumProd}(\mathbb{D}_{a,b}), \tag{5.17}$$

where $\mathbb{D}_{a,b}$ consists of pairs of vectors $(\vec{r},\vec{s})\in\mathbb{B}$ such that $r_1=s_1$, $a=r_2=s_2$ and $b=r_3=s_3$. In three dimensions, the set $\mathbb{D}_{a,b}$ is empty, since the requirements for a pair of vectors being in the set $\mathcal{D}_{a,b}$ force $\vec{r}=\vec{s}$, a contradiction.

Assuming that $d>3$, using the assumption of $\mathcal{B}(d-2)$ (in the case of $d=4$ we just apply the Littlewood–Paley inequality in the last coordinate), we see that

$$\big\|\operatorname{SumProd}(\mathbb{D}_{a,b})\big\|_{p/2} \lesssim p^{d-5/2}\cdot n^{d-7/2}. \tag{5.18}$$

Here, we have 'lost two dimensions' due to the roles of a,b. Therefore, using a trivial estimate in the parameters a,b,

$$p\left\|\left[\sum_{a,b}\operatorname{SumProd}(\mathbb{D}_{a,b})^2\right]^{1/2}\right\|_p \lesssim p^{d-3/2}n^{d-5/2}.$$

This estimate is smaller than what the other terms will give us.

Therefore, using (3.13) we can estimate

$$\big\|\operatorname{SumProd}(\mathbb{C})\big\|_p \lesssim p^{d-3/2}n^{d-5/2} + p^2\left\|\sum_{a,b}\Bigg|\sum_{\substack{(\vec{r},\vec{s})\in\mathbb{B}\\ r_2=a,\, s_3=b}} f_{\vec{r}}\cdot f_{\vec{s}}\Bigg|^2\right\|_{p/2}^{1/2}. \tag{5.19}$$

We concentrate on the latter term, and in particular expand the square,

$$\sum_{a,b}\Bigg|\sum_{\substack{(\vec{r},\vec{s})\in\mathbb{B}\\ r_2=a,\, s_3=b}} f_{\vec{r}}\cdot f_{\vec{s}}\Bigg|^2 \lesssim n^{2d-3} \tag{5.20}$$

$$+\operatorname{SumProd}\big(\mathbb{B}_1'\big)+\operatorname{SumProd}\big(\mathbb{B}_2'\big) \tag{5.21}$$

$$+\operatorname{SumProd}(\mathbb{B}''), \tag{5.22}$$

where these terms arise as follows. In forming the square on the left in (5.20), we have two pairs $(\vec{r},\vec{s}), (\vec{\underline{r}},\vec{\underline{s}})\in\mathbb{C}$ with $r_2=r_2'$ and $s_3=s_3'$. We form the product

$$f_{\vec{r}}\cdot f_{\vec{s}}\cdot f_{\vec{\underline{r}}}\cdot f_{\vec{\underline{s}}}. \tag{5.23}$$

- If the two pairs are equal, the product in (5.23) is one. There are $\lesssim n^{2d-3}$ ways to select such pairs. This is the right-hand side of (5.20).

- The collection $\mathbb{B}'_1$ consists of vectors such that $\vec{r} = \underline{\vec{r}}$ but $\vec{s} \neq \underline{\vec{s}}$, the product in (5.23) is equal to $f_{\vec{s}} \cdot f_{\underline{\vec{s}}}$ ($\mathbb{B}'_2$ is defined symmetrically). Notice that necessarily we have $s_1 = s'_1$, which is equal to r_1, and $s_3 = s'_3$. Let us set

$$\mathbb{B}'_c = \{(\vec{s}, \underline{\vec{s}}):\ s_1 = \underline{s}_1 = c;\ s_3 = \underline{s}_3\}.$$

We have 'lost' one parameter in $\mathbb{B}'_c$ and have one more coincidence, therefore, we can apply the induction hypothesis $\mathcal{B}(d-1)$ to see that

$$\left\|\operatorname{SumProd}(\mathbb{B}'_c)\right\|_p \lesssim p^{d-3/2} n^{d-5/2}.$$

It is easy to see that

$$\operatorname{SumProd}(\mathbb{B}'_1) = \sum_{\vec{r} \in \mathbb{H}_{n_1}} \operatorname{SumProd}(\mathbb{B}'_{r_1}).$$

Thus we have

$$\left\|\operatorname{SumProd}(\mathbb{B}'_1)\right\|_p \leqslant \sum_{\vec{r} \in \mathbb{H}_{n_1}} \left\|\operatorname{SumProd}(\mathbb{B}'_{r_1})\right\|_p \leqslant n^{d-1} \cdot p^{d-3/2} n^{d-5/2} = p^{d-3/2} n^{d-7/2}.$$

This controls the term in (5.21).
- The last term arises from two pairs of vectors $(\vec{r}, \vec{s}), (\underline{\vec{r}}, \vec{s}) \in \mathbb{C}$ that consist of four distinct vectors. Let us set

$$\mathbb{B}'' = \{(\vec{r}, \vec{s}, \underline{\vec{s}}, \underline{\vec{r}}):\ (\vec{r}, \vec{s}), (\underline{\vec{r}}, \underline{\vec{s}}) \in \mathbb{C},\ \vec{r} \neq \underline{\vec{r}},\ \vec{s} \neq \underline{\vec{s}}\}.$$

Here, for the sake of cleaner graphics, we have deliberately written $\vec{s}, \underline{\vec{s}}$ as the middle two vectors in the four-tuples in $\mathbb{B}''$.

It remains to bound the term in (5.22). We reduce this four-fold product back to a product of two-fold products. For integers F_1, F_2, let $\mathbb{B}''_{F_1,F_2}$ be those $(\vec{r}, \vec{s}, \underline{\vec{s}}, \underline{\vec{r}}) \in \mathbb{B}''$ with $r_1 = s_1 = F_1$ and $\underline{r}_1 = \underline{s}_1 = F_2$. Let $\mathbb{B}''_{\text{outside},F_1,F_2}$ be the projection of four-tuples in $\mathbb{B}''_{F_1,F_2}$ onto the first and fourth coordinates, and $\mathbb{B}''_{\text{inside},F_1,F_2}$ the projection onto the second and third coordinates. See Fig. 2.

For any pair $(\vec{r}, \underline{\vec{r}}) \in \mathbb{B}''_{\text{outside},F_1,F_2}$, and any two pairs

$$(\vec{s}, \underline{\vec{s}}), (\vec{\sigma}, \underline{\vec{\sigma}}) \in \mathbb{B}''_{\text{inside},F_1,F_2},$$

we have

$$(\vec{r}, \vec{s}, \underline{\vec{s}}, \underline{\vec{r}}), (\vec{r}, \vec{\sigma}, \underline{\vec{\sigma}}, \underline{\vec{r}}) \in \mathbb{B}''_{F_1,F_2}.$$

Therefore, we have the product formula

$$\operatorname{SumProd}(\mathbb{B}''_{F_1,F_2}) = \operatorname{SumProd}(\mathbb{B}''_{\text{outside},F_1,F_2}) \times \operatorname{SumProd}(\mathbb{B}''_{\text{inside},F_1,F_2}).$$

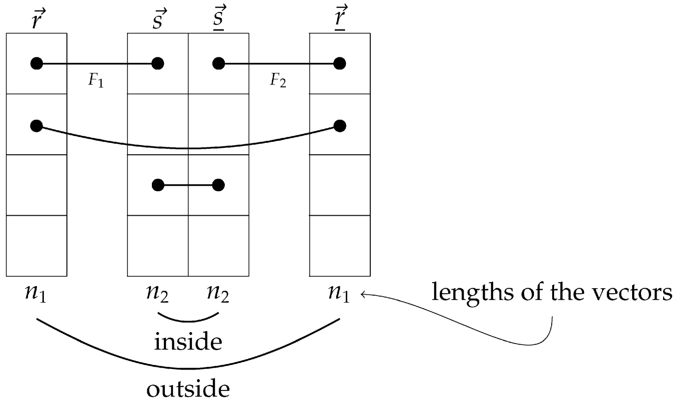

Fig. 2. The decomposition of $\mathbb{B}''_{F_1,F_2}$, in the four-dimensional case. Note that the coincidences are indicated by the connected black circles.

Notice that the pairs of vectors in $\mathbb{B}''_{\text{outside},F_1,F_2}$ have their first coordinates fixed, and have a coincidence in the second coordinate. The fixed first coordinates need not be the same, so that the lengths of the remaining coordinates are, in general, distinct. Still, we may conclude that

$$\left\| \operatorname{SumProd}\left(\mathbb{B}''_{\text{outside},F_1,F_2}\right) \right\|_p \lesssim p^{d-3/2} n^{d-5/2}.$$

This estimate is uniform in F_1, F_2. In the case of dimension $d = 3$, this follows from Lemma 5.15, while for $d > 3$ it follows from the induction hypothesis. A similar inequality holds for $\mathbb{B}''_{\text{inside},F_1,F_2}$.

Therefore, we can estimate the term in (5.22) as follows:

$$\begin{aligned}
\left\| \operatorname{SumProd}(\mathbb{B}'') \right\|_{p/2}^{1/2} &\lesssim pn \sup_{F_1,F_2} \left\| \operatorname{SumProd}\left(\mathbb{B}''_{\text{outside},F_1,F_2}\right) \times \operatorname{SumProd}\left(\mathbb{B}''_{\text{inside},F_1,F_2}\right) \right\|_{p/2}^{1/2} \\
&\lesssim pn \sup_{F_1,F_2} \left\| \operatorname{SumProd}\left(\mathbb{B}''_{\text{outside},F_1,F_2}\right) \right\|_p^{1/2} \times \left\| \operatorname{SumProd}\left(\mathbb{B}''_{\text{inside},F_1,F_2}\right) \right\|_p^{1/2} \\
&\lesssim (pn)^{d-3/2}.
\end{aligned}$$

Our proof is complete. Assuming $\mathcal{B}(d-1)$, $d \geqslant 4$, we have proved $\mathcal{C}(d)$. We have also proved $\mathcal{C}(3)$. The fact that $\mathcal{B}(3)$ holds follows from the argument below.

Assuming $\mathcal{C}(d)$ and $\mathcal{D}(d)$, we prove $\mathcal{B}(d)$. Fix $p, n \geqslant 3$, a collection of vectors $\mathbb{B}$ and r functions which achieve the supremum in (5.8). Write this collection as

$$\mathbb{B} = \mathbb{D} \cup \bigcup_{2 \leqslant i \neq j \leqslant d} \mathbb{C}_{i,j},$$

where $\mathbb{C}_{i,j}$ consists of those pairs $(\vec{r}, \vec{s}) \in \mathbb{B}$ such that i is the first coordinate for which $r_i > s_i$ and j is the first coordinate for which $r_j < s_j$. Then, the collections $\mathbb{C}_{i,j}$ are pairwise disjoint, and the collection $\mathbb{D}$ consists of all pairs not in some $\mathbb{C}_{i,j}$. Thus,

$$\operatorname{SumProd}(\mathbb{B}) = \operatorname{SumProd}(\mathbb{D}) + \sum_{2 \leqslant i \neq j \leqslant d} \operatorname{SumProd}(\mathbb{C}_{i,j}).$$

After a harmless permutation of indices, the inequalities $\mathcal{C}(d)$ apply to the collections $\mathbb{C}_{i,j}$. The (unconditional) inequalities $\mathcal{D}$ apply to the collection $\mathbb{D}$. The proof is complete. □

6. Corollaries of the Beck gain

Theorem 3.10 implies an exponential estimate of order $\exp(L^{2/(d-1)})$ for sums of $\vec{r}$ functions. In fact, we can derive a *subgaussian* estimate for such sums, for moderate deviations, and moreover, in order to have a gain of order n^{c/d^2} in our main theorem, we need to use this estimate.

Theorem 6.1. *Using the notation of* (4.2) *and* (4.3)*, we have this estimate, valid for all* $1 \leqslant t \leqslant q$*,*

$$\|\rho F_t\|_p \lesssim \sqrt{p}, \quad 1 \leqslant p \leqslant cn^{\frac{1-2\varepsilon}{2d-1}}. \tag{6.2}$$

As a consequence, we have the distributional estimate

$$\mathbb{P}\big(|\rho F_t| > x\big) \lesssim \exp\big(-cx^2\big), \quad x < cn^{\frac{1-2\varepsilon}{4d-2}}. \tag{6.3}$$

Here $0 < c < 1$ *is an absolute constant.*

To use (6.3), we need $q^b = a^b n^{\epsilon \cdot b} < cn^{\frac{1}{4d-6}}$, and so $\epsilon \simeq 1/d$ is the optimal value for ϵ that this proof will give.

Proof. Recall that

$$F_t = \sum_{\vec{r} \in \mathbb{A}_t} f_{\vec{r}},$$

where $\mathbb{A}_t := \{\vec{r} \in \mathbb{H}_n\colon\ r_1 \in I_t\}$, and I_t is an interval of integers of length n/q, so that $\sharp \mathbb{A}_t \simeq n^{d-1}/q \simeq \rho^{-2}$.

Apply the Littlewood–Paley inequality in the first coordinate. This results in the estimate

$$\begin{aligned}
\|\rho F_t\|_p &\lesssim \sqrt{p} \left\| \left[\sum_{s \in I_j} \left| \rho \sum_{\vec{r}\colon\, r_1 = s} f_{\vec{r}} \right|^2 \right]^{1/2} \right\|_p \\
&\lesssim \sqrt{p} \left\| 1 + \rho^2 \Phi_{t,t,1} \right\|_{p/2}^{1/2} \\
&\lesssim \sqrt{p} \left\{ 1 + \left\| \rho^2 \Phi_{t,t,1} \right\|_{p/2}^{1/2} \right\},
\end{aligned}$$

where $\Phi_{t,t,1}$ is defined in (5.1). Here it is important to use the constants in the Littlewood–Paley inequalities that give the correct order of growth of $\sqrt{p}$. Of course the terms $\Phi_{t,t,1}$ are controlled by the estimate in (5.3). In particular, we have

$$\left\| \rho^2 \Phi_{t,t,1} \right\|_p \lesssim \frac{q}{n^{d-1}} p^{d-1/2} n^{d-3/2} \lesssim q p^{d-1/2} n^{-1/2} \lesssim 1. \tag{6.4}$$

Hence (6.2) follows.

The second distributional inequality is a well-known consequence of the norm inequality. Namely, one has the inequality below, valid for all x:

$$\mathbb{P}(\rho F_t > x) \leqslant C^p p^{p/2} x^{-p}, \quad 1 \leqslant p \leqslant c n^{\frac{1-2\varepsilon}{2d-1}}.$$

If x is as in (6.3), we can take $p \simeq x^2$ to prove the claimed exponential squared bound. □

We shall now use the Beck gain to prove the crucial L^2 estimate (4.10) of Lemma 4.8. We actually need a slightly more general inequality:

Lemma 6.5. *We have the following estimate*:

$$\sup_{V \subset \{1,\dots,q\}} \mathbb{E} \prod_{v \in V} (1 + \tilde{\rho} F_t)^2 \lesssim \exp\bigl(a' q^{2b}\bigr). \tag{6.6}$$

The supremum over V will be an immediate consequence of the proof below, and so we do not address it specifically.

Proof of (4.10). Let us give the essential initial observation. We expand

$$\mathbb{E} \prod_{j=1}^{q} (1 + \tilde{\rho} F_j)^2 = \mathbb{E} \prod_{j=1}^{q} \bigl(1 + 2\tilde{\rho} F_j + (\tilde{\rho} F_j)^2\bigr).$$

Hold the last $d-1$ coordinates, $x_2, \dots, x_d$, fixed and let $\mathcal{F}$ be the sigma field generated by $F_1, \dots, F_{q-1}$. We have

$$\begin{aligned} \mathbb{E}\bigl(1 + 2\tilde{\rho} F_q + (\tilde{\rho} F_q)^2 \bigm| \mathcal{F}\bigr) &= 1 + \mathbb{E}\bigl((\tilde{\rho} F_q)^2 \bigm| \mathcal{F}\bigr) \\ &= 1 + a^2 q^{2b-1} + \tilde{\rho}^2 \mathbb{E}(\Phi_{q,q,1} | \mathcal{F}), \end{aligned} \tag{6.7}$$

where $\Phi_{q,q,1}$ is defined in (5.1). Then, we see that

$$\begin{aligned} \mathbb{E} \prod_{v=1}^{q} \bigl(1 + 2\tilde{\rho} F_t + (\tilde{\rho} F_t)^2\bigr) &= \mathbb{E}\left\{ \prod_{v=1}^{q-1} \bigl(1 + 2\tilde{\rho} F_t + (\tilde{\rho} F_t)^2\bigr) \times \mathbb{E}\bigl(1 + 2\tilde{\rho} F_t + (\tilde{\rho} F_t)^2 \bigm| \mathcal{F}\bigr) \right\} \\ &\leqslant \bigl(1 + a^2 q^{2b-1}\bigr) \mathbb{E} \prod_{v=1}^{q-1} \bigl(1 + 2\tilde{\rho} F_t + (\tilde{\rho} F_t)^2\bigr) \qquad (6.8) \\ &\quad + \mathbb{E} \bigl| \tilde{\rho}^2 \Phi_{q,q,1} \bigr| \cdot \prod_{v=1}^{q-1} \bigl(1 + 2\tilde{\rho} F_t + (\tilde{\rho} F_t)^2\bigr). \qquad (6.9) \end{aligned}$$

This is the main observation: one should induct on (6.8), while treating the term in (6.9) as an error, as the Beck gain estimate (5.3) applies to it.

Let us set up notation to implement this line of approach. Set

$$N(V;r) := \left\| \prod_{t=1}^{V} (1 + \tilde{\rho} F_t) \right\|_r, \quad V = 1, \ldots, q.$$

We will obtain a very crude estimate for these numbers for $r = 4$. Fortunately, this is relatively easy for us to obtain. Namely, q is small enough that we can use the inequalities (6.2) to see that

$$\begin{aligned} N(V;4) &\leqslant \prod_{v=1}^{V} \|1 + \tilde{\rho} F_t\|_{4V} \\ &\leqslant \left(1 + Cq^{1/2+b}\right)^V \\ &\leqslant (Cq)^q. \end{aligned}$$

We have the estimate below from Hölder's inequality

$$N\left(V; 2(1-1/q)^{-1}\right) \leqslant N(V;2)^{1-1/q} \cdot N(V;4)^{1/q}. \tag{6.10}$$

We see that (6.8)–(6.10) give us the inequality

$$\begin{aligned} N(V+1;2)^2 &\leqslant \left(1 + a^2 q^{2b-1}\right) N(V;2)^2 + C \cdot N\left(V; 2(1-1/q)^{-1}\right)^2 \cdot \left\| \tilde{\rho}^2 \Phi_{V+1,V+1,1} \right\|_q \\ &\leqslant \left(1 + a^2 q^{2b-1}\right) N(V;2)^2 + C N(V;2)^{2-2/q} \cdot N(V;4)^{2/q} \left\| \tilde{\rho}^2 \Phi_{V+1,V+1,1} \right\|_q \\ &\leqslant \left(1 + a^2 q^{2b-1}\right) N(V;2)^2 + C q^{d+2} n^{-1/2} N(V;2)^{2-2/q}. \end{aligned} \tag{6.11}$$

In the last line we have used the inequality (5.3). Of course we only apply this as long as $N(V;2) \geqslant 1$. Assuming this is true for all $V \geqslant 1$, we see that

$$N(V+1;2)^2 \leqslant \left(1 + a^2 q^{2b-1} + C q^{d+2} n^{-1/2}\right) N(V;2)^2.$$

And so, by induction,

$$N(q;2) \lesssim \left(1 + a^2 q^{2b-1} + C q^{d+2} n^{-1/2}\right)^{q/2} \lesssim \mathrm{e}^{2aq^{2b}}.$$

Here, the last inequality will be true for large n, provided that ε in the definition of q (4.1) is small. Indeed, we need

$$a^2 q^{2b-1} \geqslant C q^{d+2} n^{-1/2}.$$

Or equivalently,

$$a^2 n^{1/2} \gtrsim q^{d+5/2}.$$

Comparing to the definition of q in (4.1), we see that the proof is finished. □

One should notice that the results of this section suggest that our methods give a gain of the order $\frac{1}{d}$.

7. The Beck gain with fixed parameters

We will need to analyze longer products of r functions. These longer products will be reduced to the case of a slightly more general version of the Beck gain Lemma 5.2. Namely, we will consider sums of products of two r functions, but impose the additional restriction for some coordinates in the pair of vectors to have fixed values. Let $\vec{a} \in \mathbb{N}^{F_1}$ and $\vec{b} \in \mathbb{N}^{F_2}$ be integer vectors with lengths $|\vec{a}|, |\vec{b}| < n$. We will be estimating the quantity:

$$\mathcal{B}(F_1, F_2) = \sup_{\vec{a},\vec{b}, j_1<j_2} \sup_{\mathbb{B}} \big\| \operatorname{SumProd}(\mathbb{B}) \big\|_p, \quad d, n, p \geqslant 3. \tag{7.1}$$

The inner supremum is formed over all $\mathbb{B} \subset \mathbb{H}_n \times \mathbb{H}_n$ and all r functions subject to these conditions:

- $\vec{r} \in \mathbb{A}_{j_1}$, $\vec{s} \in \mathbb{A}_{j_2}$, where $j_1 < j_2$ (i.e. s_1 is the maximum in the first coordinate).
- There is a coincidence in the second coordinate: For all $(\vec{r}, \vec{s}) \in \mathbb{B}$, we have $\vec{r} \neq \vec{s}$ and $r_2 = s_2$.
- For $k = 1, \dots, F_1$, we have $r_{k+2} = a_k$. (F_1 coordinates of $\vec{r}$ are fixed.)
- For $k = 1, \dots, F_2$, we have $s_{F_1+k+2} = b_k$. (F_2 coordinates of $\vec{s}$ are fixed, and these coordinates are distinct from the other vector.)

We have the following estimate, which gives an average Beck gain of $n^{1/8}$ for each of the two functions in the product.

Lemma 7.2. *We have the inequality below valid for all dimensions $d \geqslant 3$,*

$$\mathcal{B}(F_1, F_2) \lesssim p^{d-1-\frac{F_1+F_2}{2}-\frac{1}{4}} n^{d-1-\frac{F_1+F_2}{2}-\frac{1}{4}}, \quad p, n \geqslant 3.$$

Proof. We will reduce this situation to the Beck gain proven before. Let $\mathbb{B}$ be as above. First of all, we shall apply the Littlewood–Paley inequality in the first coordinate. Notice that the maximum in this coordinate is automatically s_1.

$$\big\| \operatorname{SumProd}(\mathbb{B}) \big\|_p \lesssim \sqrt{p} \Bigg\| \sum_{c \in I_{j_2}} \Bigg| \sum_{\substack{(\vec{r},\vec{s}) \in \mathbb{B} \\ s_1 = c}} f_{\vec{r}} \cdot f_{\vec{s}} \Bigg|^2 \Bigg\|_{p/2}^{1/2}. \tag{7.3}$$

We concentrate on the latter term, and in particular expand the square,

$$\sqrt{p} \Bigg\| \sum_{c \in \mathbb{I}_{j_2}} \Bigg| \sum_{\substack{(\vec{r},\vec{s}) \in \mathbb{B} \\ s_1 = c}} f_{\vec{r}} \cdot f_{\vec{s}} \Bigg|^2 \Bigg\|_{p/2}^{1/2} = \sqrt{p} \Bigg\| \sum_{\substack{(\vec{r},\vec{s},\underline{\vec{r}},\underline{\vec{s}}) \in \mathbb{B}\times\mathbb{B} \\ s_1 = \underline{s}_1}} f_{\vec{r}} \cdot f_{\vec{s}} \cdot f_{\underline{\vec{r}}} \cdot f_{\underline{\vec{s}}} \Bigg\|_{p/2}^{1/2} \tag{7.4}$$

$$\leqslant \sqrt{p} n \max_{c \neq \underline{c}} \Bigg\| \sum_{\substack{(\vec{r},\vec{s},\underline{\vec{r}},\underline{\vec{s}}) \in \mathbb{B}\times\mathbb{B} \\ s_1 = \underline{s}_1;\, r_2 = s_2 = c;\, \underline{r}_2 = \underline{s}_2 = \underline{c}}} f_{\vec{r}} \cdot f_{\vec{s}} \cdot f_{\underline{\vec{r}}} \cdot f_{\underline{\vec{s}}} \Bigg\|_{p/2}^{1/2} \tag{7.5}$$

$$+ \sqrt{p}\sqrt{n} \max_{c} \Bigg\| \sum_{\substack{(\vec{r},\vec{s},\underline{\vec{r}},\underline{\vec{s}}) \in \mathbb{B}\times\mathbb{B} \\ s_1 = \underline{s}_1;\, r_2 = s_2 = \underline{r}_2 = \underline{s}_2 = c}} f_{\vec{r}} \cdot f_{\vec{s}} \cdot f_{\underline{\vec{r}}} \cdot f_{\underline{\vec{s}}} \Bigg\|_{p/2}^{1/2}. \tag{7.6}$$

We start with the estimates for the first term above (7.5):

$$\begin{aligned}
&\sqrt{pn}\max_{c\neq \underline{c}}\Bigg\| \sum_{\substack{(\vec r,\vec s,\vec{\underline{r}},\vec{\underline{s}})\in\mathbb{B}\times\mathbb{B}\\ s_1=\underline{s}_1;\, r_2=s_2=c;\, \underline{r}_2=\underline{s}_2=\underline{c}}} f_{\vec r}\cdot f_{\vec s}\cdot f_{\vec{\underline{r}}}\cdot f_{\vec{\underline{s}}}\Bigg\|_{p/2}^{1/2}\\
&\quad=\sqrt{pn}\max_{c\neq \underline{c}}\Bigg\|\bigg(\sum_{(\vec r,\vec{\underline{r}})\in\mathbb{B}_1} f_{\vec r}\cdot f_{\vec{\underline{r}}}\bigg)\times\bigg(\sum_{(\vec s,\vec{\underline{s}})\in\mathbb{B}_2} f_{\vec s}\cdot f_{\vec{\underline{s}}}\bigg)\Bigg\|_{p/2}^{1/2}\\
&\quad\leqslant\sqrt{pn}\max_{c\neq \underline{c}}\Bigg\|\sum_{(\vec r,\vec{\underline{r}})\in\mathbb{B}_1} f_{\vec r}\cdot f_{\vec{\underline{r}}}\Bigg\|_p^{1/2}\Bigg\|\sum_{(\vec s,\vec{\underline{s}})\in\mathbb{B}_2} f_{\vec s}\cdot f_{\vec{\underline{s}}}\Bigg\|_p^{1/2}.
\end{aligned}$$

Here $\mathbb{B}_1$ is defined to consist of pairs $(\vec r,\vec{\underline{r}})\in\mathbb{A}_{j_1}^2$ which satisfy the following:

- For $k=1,\dots,F_1$, we have $r_{k+2}=\underline{r}_{k+2}=a_k$.
- $r_2=c$, $\underline{r}_2=\underline{c}$.

And similarly $\mathbb{B}_2$ consists of pairs $(\vec s,\vec{\underline{s}})\in\mathbb{A}_{j_2}^2$ with the properties:

- For $k=1,\dots,F_2$, we have $s_{k+F_1+2}=\underline{s}_{k+F_1+2}=b_k$.
- $s_2=c$, $\underline{s}_2=\underline{c}$.
- Moreover, we have $s_1=\underline{s}_1$.

Notice that because of the last condition and the fact that $c\neq\underline{c}$ (i.e., $\vec s\neq\vec{\underline{s}}$), the Beck gain (Lemma 5.2) applies to this family of pairs, giving a gain of $n^{1/2}$, while $\mathbb{B}_1$ will be estimated by simple parameter counting, supplying no gain. We have

$$\begin{aligned}
&\big\|\mathrm{SumProd}(\mathbb{B}_1)\big\|_p\lesssim (pn)^{d-2-F_1},\\
&\big\|\mathrm{SumProd}(\mathbb{B}_2)\big\|_p\lesssim p^{d-3/2-F_2}n^{d-2-F_2-1/2}.
\end{aligned}$$

And thus we can estimate the term (7.5) by

$$\begin{aligned}
&\sqrt{pn}\Bigg\| \sum_{\substack{(\vec r,\vec s,\vec{\underline{r}},\vec{\underline{s}})\in\mathbb{B}\times\mathbb{B}\\ s_1=\underline{s}_1;\, r_2=s_2=c;\, \underline{r}_2=\underline{s}_2=\underline{c}}} f_{\vec r}\cdot f_{\vec s}\cdot f_{\vec{\underline{r}}}\cdot f_{\vec{\underline{s}}}\Bigg\|_{p/2}^{1/2}\\
&\quad\lesssim\sqrt{pn}\big((pn)^{d-2-F_1}\big)^{1/2}\big(p^{d-3/2-F_2}n^{d-2-F_2-1/2}\big)^{1/2}\\
&\quad=(pn)^{d-1-\frac{F_1+F_2}{2}-\frac14}.
\end{aligned}$$

The second term (7.6) satisfies the same bound in n. This can be shown by simple parameter counting, the gain comes from the loss of one parameter since $c=\underline{c}$.

We remark that in this version of the Beck gain 'error terms' do not arise, since we apply Littlewood–Paley inequality only in the first coordinate, where we already have a natural order. Thus we do not need to use the conditional expectation argument as in the proof of Lemma 5.2. □

8. The Beck gain for longer coincidences

In the present section we treat longer coincidences. This requires a careful analysis of the variety of ways that a product can fail to be strongly distinct. That is, we need to understand the variety of ways that coincidences can arise, and how coincidences can contribute to a smaller norm. Following Beck, we will use the language of graph theory to describe these general patterns of coincidences.

8.1. Graph theory nomenclature

We adopt familiar nomenclature from graph theory, although there is no graph theoretical fact that we need, rather the use of this language is just a convenient way to do some bookkeeping. The class of graphs that we are interested in satisfies particular properties. A $d-1$ *colored graph* G is the tuple $(V(G), E_2, E_3, \dots, E_d)$, of the *vertex set* $V(G) \subset \{1, \dots, q\}$, and *edge sets* $E_2, E_3, \dots, E_d$*, of colors* $2, 3, \dots, d$*, respectively*. Edge sets are subsets of

$$E_j \subset V(G) \times V(G) - \{(k,k) \mid k \in V(G)\}.$$

Edges are symmetric, thus if $(v, v') \in E_j$ then necessarily $(v', v) \in E_j$.

A *clique of color* j is a maximal subset $Q \subset V(G)$ such that for all $v \neq v' \in Q$ we have $(v, v') \in E_j$. By *maximality*, we mean that no strictly larger set of vertices $Q' \supset Q$ satisfies this condition.

Call a graph G *admissible* iff:

- The edges sets, in all $d-1$ colors, decompose into a union of cliques.
- If Q_k's are cliques of color k ($k = 2, \dots, d$), then $\bigcap_{k=2}^{d} Q_k$ contains at most one vertex.
- Every vertex is in at least one clique.

A graph G is *connected* iff for any two vertices in the graph, there is a path that connects them. A *path* in the graph G is a sequence of vertices $v_1, \dots, v_k$ with an edge of *any color,* spanning adjacent vertices , that is $(v_j, v_{j+1}) \in \bigcup_{k=2}^{d} E_k$.

8.2. Reduction to admissible graphs

It is clear that admissible graphs as defined above are naturally associated to sums of products of r functions. Given admissible graph G on vertices V, we set $X(G)$ to be those tuples of r vectors

$$\vec{r}_v \in \prod_{v \in V} \mathbb{A}_v$$

so that if (v, v') is an edge of color j in G, then $r_{v,j} = r_{v',j}$.

We shall introduce the following counting parameter: for an admissible graph G, its index, $\operatorname{ind}(G)$, is defined as

$$\operatorname{ind}(G) = \sum_{Q \text{ is a clique}} (\sharp Q - 1). \tag{8.1}$$

Effectively, the index of G is the least number of equalities, needed to define $X(G)$, in other words, the number of coincidences. In particular, for the graphs, corresponding to the simplest case of the Beck gain, the index is one.

With these definitions at hand, it is not hard to obtain the inclusion–exclusion formula, relating admissible graphs and the 'not strongly distinct' part of the Riesz product:

$$\Psi^{\neg} = \sum_{G \text{ admissible}} (-1)^{\operatorname{ind}(G)+1} \tilde{\rho}^{|V(G)|} \operatorname{SumProd}\bigl(X(G)\bigr) \cdot \prod_{t \notin V(G)} (1 + \tilde{\rho} F_t). \tag{8.2}$$

We will prove the following theorem:

Theorem 8.3 *(Beck gain for graphs). For an admissible graph G on vertices V we have the estimate below for positive, finite constants C_0, C_1, C_2, C_3:*

$$\rho^{|V|} \bigl\| \operatorname{SumProd}\bigl(X(G)\bigr)\bigr\|_p \leqslant \bigl[C_0 |V|^{C_1} p^{C_2} q^{C_3} n^{-\eta}\bigr]^{|V|}, \quad 2 < p < \infty. \tag{8.4}$$

The most significant term on the right is $n^{-\eta}$. It shows that as the number of coincidences goes up, the corresponding 'Beck gain' improves. Notice that for the other terms on the right, C_0 is a constant; $|V| \leqslant q \leqslant n^{\epsilon}$, where we can choose $0 < \epsilon$ as a function of η; and while the inequality above holds for all $2 \leqslant p < \infty$, we will only need to apply it for $p \lesssim q^{2b} \leqslant n^{\epsilon/2}$. That is, the $n^{-\eta}$ is the dominant term on the right. This theorem, together with the fact that there are at most $|V|^{2d|V|}$ admissible graphs on the vertex set V, yields the boundedness of the sum in (4.16).

8.3. Norm estimates for admissible graphs

We begin the proof of Theorem 8.3 with a further reduction to connected admissible graphs. Let us write $G \in \mathrm{BG}(C_0, C_1, C_2, C_3, \eta)$ if the estimates (8.4) hold. ('BG' for 'Beck gain.') We need to see that all admissible graphs are in $\mathrm{BG}(C_0, C_1, C_2, C_3, \eta)$ for non-negative, finite choices of the relevant constants.

Lemma 8.5. *Let C_0, C_1, C_2, C_3, η be non-negative constants. Suppose that G is an admissible graph, and that it can be written as a union of subgraphs $G_1, \ldots, G_k$ on disjoint vertex sets, where all $G_j \in \mathrm{BG}(C_0, C_1, C_2, C_3, \eta)$. Then,*

$$G \in \mathrm{BG}(C_0, C_1, C_2, C_2 + C_3, \eta).$$

With this lemma, we will identify a small class of graphs for which we can verify the property (8.4) directly, and then appeal to this lemma to deduce Theorem 8.3. Accordingly, we modify our notation. If $\mathcal{G}$ is a class of graphs, we write $\mathcal{G} \subset \mathrm{BG}(\eta)$ if there are constants C_0, C_1, C_2, C_3 such that $\mathcal{G} \subset \mathrm{BG}(C_0, C_1, C_2, C_3, \eta)$.

Proof of Lemma 8.5. We then have by Proposition 8.6

$$\operatorname{SumProd}\bigl(X(G)\bigr) = \prod_{j=1}^{k} \operatorname{SumProd}\bigl(X(G_j)\bigr).$$

Using Hölder's inequality, we can estimate

$$\begin{aligned}\rho^{|V|}\big\|\operatorname{SumProd}\big(X(G)\big)\big\|_p &\leqslant \prod_{j=1}^{k}\rho^{|V_j|}\big\|\operatorname{SumProd}\big(X(G_j)\big)\big\|_{kp}\\ &\leqslant \prod_{j=1}^{k}\big[C_0(kp)^{C_1}q^{C_2}n^{-\eta}\big]^{|V_j|}\\ &\leqslant \big[C_0p^{C_1}q^{C_2+C_1}n^{-\eta}\big]^{|V|}.\end{aligned}$$

Here, we use the fact that since the graphs are non-empty, we necessarily have $k \leqslant q$. $\square$

Proposition 8.6. *Let $G_1, \ldots, G_p$ be admissible graphs on pairwise disjoint vertex sets $V_1, \ldots, V_p$. Extend these graphs in the natural way to a graph G on the vertex set $V = \bigcup V_t$. Then, we have*

$$\operatorname{SumProd}\big(X(G)\big) = \prod_{t=1}^{p}\operatorname{SumProd}\big(X(G_t)\big).$$

8.4. Connected graphs have the Beck gain

We single out for special consideration the connected admissible graphs G . Let $\mathcal{G}_{\text{connected}}$ be the collection of all admissible connected graphs on $V \subset \{1, \ldots, q\}$.

Lemma 8.7. *We have $\mathcal{G}_{\text{connected}} \subset \mathrm{BG}(\eta)$ for some $\eta > 0$.*

The point of this proof is that we will reduce this question to a much simpler key fact, namely Lemma 7.2, which we restate here in our current notation.[3]

Let $\mathcal{G}_{\text{fixed}}(2)$ be the set of graphs—and sets of r functions associated with the graphs—with these properties:

- G is a connected graph on two vertices $\{v, v'\}$. That is, there is at least one edge that connects these two vertices. Denote by $C \subset \{2, \ldots, d\}$ the set of coordinates corresponding to the edges.
- There are a set of coordinates $F_v, F_{v'} \subset \{2, \ldots, d\}$ that are disjoint from the set of edges, and two vectors $\vec{a} \in \mathbb{N}^{F_v}$ and $\vec{a}' \in \mathbb{N}^{F_{v'}}$, so that we define

$$Y(G) := \big\{(\vec{r}_v, \vec{r}_{v'}) \in \mathbb{H}_n\colon\ r_{v,j} = r_{v',j}\ \forall j \in C;\ r_{v,k} = a_k\ \forall k \in F_v;\ r_{v',k} = a_k\ \forall k \in F_{v'}\big\}.$$

These are in essence the assumptions of Lemma 7.2. This lemma proves that

$$\big\|\operatorname{SumProd}\big(Y(G)\big)\big\|_p \lesssim p^d n^\sigma, \quad \sigma = d - 1 - \frac{F_v + F_{v'}}{2} - \frac{1}{4}.$$

By abuse of notation, let us summarize this inequality by the inclusion $\mathcal{G}_{\text{fixed}}(2) \subset \mathrm{BG}(C_0, C_1, d/2, 0, 1/8)$. Or, even more briefly, as $\mathcal{G}_{\text{fixed}}(2) \subset \mathrm{BG}(1/8)$. That is, there is a gain

[3] The only points that recommend the proof we describe here is that it is easy to state and delivers a gain. Clearly, a more sustained analysis, yielding a larger gain would result in an improved result on the Small Ball Conjecture.

of $\frac{1}{8}$ for each vertex. It follows from the proof of Lemma 8.5, that if G is any graph whose connected components are each elements of $\mathcal{G}_{\text{fixed}}(2)$, then $G \in \mathrm{BG}(1/8)$.

Our line of attack on this lemma is to take a general connected graph G, use the triangle inequality to assign fixed values to a number of edges, making the connected components of the new graph to be elements of $\mathcal{G}_{\text{fixed}}(2)$. The proportion of vertices that will be in one of these graphs will be at least $1/2d$ of all vertices. And therefore connected graphs will be in $\mathrm{BG}(1/16d)$.

Remark 8.8. A heuristic guides this argument. The normalization $\rho^{|V|}$ in (8.4) assigns a weight $n^{-1/2}$ to each free parameter of $X(G)$, ignoring losses of parameters from the edges of G. If (v, v') is an edge in the graph, and we assign the edge one of n possible values, the full power of n is exactly compensated by the collective weight of the two parameters in the edge. Therefore, we are free to fix a fixed proportion of edges in the graph, obtaining a Beck gain on the remaining proportion. In this argument, if the edge is in a clique of size at least $k \geqslant 3$, specifying a single value on this clique actually leads to a positive gain of $n^{-k/2+1}$. In other words, graphs, all of whose cliques are of size two, are extremal with respect to this analysis (see Lemma 8.9). This heuristic is made precise in the proof below.

By 'deleting a clique' we shall mean fixing a value of the coincidence which corresponds to that clique. Let $G \in \mathcal{G}_{\text{connected}}$. Following the heuristic above, in the first step of the algorithm we delete all cliques of size at least 3 in G.

After this step G breaks down into connected components, which are admissible graphs with cliques only of size 2 (and, possibly, some singletons). Next, we want to obtain an estimate for such graphs.

Lemma 8.9. *Suppose $\widetilde{G} \in \mathcal{G}_{\text{connected}}$ has cliques of size at most 2. Then $\widetilde{G} \in \mathrm{BG}(\frac{1}{16d})$.*

To prove this statement we shall use the following property of $\widetilde{G}$:

- The degree of each vertex in $\widetilde{G}$ is at most $d-1$ (since the degree in each color is at most one).

Let $\widetilde{V}$ be the set of vertices of $\widetilde{G}$, and $\widetilde{E}$ be the set of all its edges. The point is to select a maximal subset $\widetilde{E}_{\text{indpndt}}$ of *independent* edges. That is, no two edges in $\widetilde{E}_{\text{indpndt}}$, regardless of color, have a common vertex. It is an elementary fact that we can take

$$|\widetilde{E}_{\text{indpndt}}| \geqslant \tfrac{1}{2d-3}|\widetilde{E}|. \tag{8.10}$$

Indeed, each edge in $\widetilde{G}$ shares a vertex with at most $2d-4$ distinct edges, which observation directly implies the inequality above.

We delete all other edges of $\widetilde{G}$ (i.e. we fix some choice of parameters for the corresponding coincidences) and thus $\widetilde{G}$ breaks down into a number of components each of which is either a singleton or a graph with two vertices and one edge. The latter components correspond exactly to the situation in which the Beck gain of the previous section is applicable. Let us denote these pairs by $G_k \in \mathcal{G}_{\text{fixed}}(2)$, $k = 1, \ldots, N = |\widetilde{E}_{\text{indpndt}}|$; the singletons – by v_j, $j = 1, \ldots, |\widetilde{V}| - 2N$. Let also $E' = \widetilde{E} - \widetilde{E}_{\text{indpndnt}}$ denote the set of all deleted edges in $\widetilde{G}$. Denote also by F_k the number of fixed parameters in $X(G_k)$ and F'_j will be the number of fixed parameters in $\vec{r}_{v_j}$. We have the following relations:

$$2|E'| = 2|\widetilde{E} - \widetilde{E}_{\text{indpndnt}}| = \sum_{k=1}^{N} F_k + \sum_{j=1}^{|\widetilde{V}|-2N} F'_j, \tag{8.11}$$

and, since $\widetilde{G}$ is connected, it has at least $|V(G)| - 1$ edges, thus

$$N \geqslant \frac{|\widetilde{E}|}{2d-3} \geqslant \frac{|\widetilde{V}|-1}{2d-3} \geqslant \frac{|\widetilde{V}|}{2(2d-3)} \geqslant \frac{|\widetilde{V}|}{4d}. \tag{8.12}$$

Besides, by Proposition 8.6, we obtain the following equality (the sum below is taken over all choices of parameters on the 'deleted' edges):

$$\operatorname{SumProd}\bigl(X(\widetilde{G})\bigr) = \sum \prod_{k=1}^{N} \operatorname{SumProd}\bigl(X(G_k)\bigr) \cdot \prod_{j=1}^{|\widetilde{V}|-2N} \operatorname{SumProd}\bigl(X(v_j)\bigr). \tag{8.13}$$

Now we apply the triangle inequality, Hölder's inequality, the relations (8.11) and (8.12), and the Beck gain in the form of Lemma 7.2 to estimate ($\kappa = |\widetilde{V}| - N < q$):

$$\begin{aligned}
\rho^{|\widetilde{V}|} \bigl\| \operatorname{SumProd}\bigl(X(\widetilde{G})\bigr) \bigr\|_p &\leqslant n^{|E'|} \cdot \prod_{k=1}^{N} \rho^2 \bigl\| \operatorname{SumProd}\bigl(X(G_k)\bigr) \bigr\|_{\kappa p} \cdot \prod_{j=1}^{|\widetilde{V}|-2N} \rho \| f_{\vec{r}_{v_j}} \|_{\kappa p} \\
&\lesssim n^{|E'|} \cdot \prod_{k=1}^{N} \bigl[\rho^2 (\kappa p n)^{d-1-\frac{F_k}{2}-\frac{1}{4}}\bigr] \cdot \prod_{j=1}^{|\widetilde{V}|-2N} \bigl[\rho (\kappa p n)^{\frac{d-1}{2}-\frac{F'_j}{2}}\bigr] \\
&\lesssim \bigl[C p^{\frac{d-1}{2}} q^{\frac{d}{2}}\bigr]^{|\widetilde{V}|} \cdot n^{-\frac{N}{4}} \lesssim \bigl[p^{\frac{d-1}{2}} q^{\frac{d}{2}} n^{-\frac{1}{16d}}\bigr]^{|\widetilde{V}|}.
\end{aligned}$$

This proves Lemma 8.9. The point of passing to the collection of independent edges is that $\operatorname{SumProd}(X(\widetilde{G}))$ splits into a product of terms associated with graphs in $\mathcal{G}_{\text{fixed}}(2)$. Each of these graphs leads to a gain of at least $\frac{1}{8}$ for each vertex. But by (8.11), there are at least $\frac{1}{2d}|V(G)|$ vertices for which we will get this gain. This shows that $G \in \mathrm{BG}(1/16d)$.

We can now proceed to prove Lemma 8.7—the proof will be in the same spirit. After we delete "large" (of size at least 3) cliques of G, this graph decomposed into some singletons and components as in Lemma 8.9 (but with some parameters fixed). Denote these components by $\widetilde{G}_k$, $k = 1, \ldots, n_1$, and the singletons by u_j, $j = 1, \ldots, n_2$. Let f_k be the number of fixed parameters in $X(\widetilde{G}_k)$ and f'_j—the number of fixed parameters in $\vec{r}_{u_j}$. Notice that the proof of Lemma 8.9 can be trivially adapted to the case when some parameters are fixed to obtain the estimate

$$\rho^{|\widetilde{V}_k|} \bigl\| \operatorname{SumProd}\bigl(X(\widetilde{G}_k)\bigr) \bigr\|_p \leqslant \bigl[C p^{\frac{d-1}{2}} q^{\frac{d}{2}} n^{-\frac{1}{16d}}\bigr]^{|\widetilde{V}_k|} n^{-\frac{f_k}{2}}. \tag{8.14}$$

Also, if we denote by K the total number of fixed cliques, one can see that, since all the cliques had size at least 3, we have the inequality

$$3K \leqslant \sum_{k=1}^{n_1} f_k + \sum_{j=1}^{n_2} f'_j. \tag{8.15}$$

Let us write the set of vertices of G as $V = V_1 \cup V_2$, where V_1 are the vertices involved in at least one of the deleted cliques and V_2 are all the other vertices. It is easy to see that $V_2 \subset \bigcup_{k=1}^{n_1} V(\widetilde{G}_k)$. Indeed, all the vertices that became singletons had to be a part of one of the deleted cliques. Thus,

$$|V_2| \leqslant \sum_{k=1}^{n_1} \bigl|V(\widetilde{G}_k)\bigr|. \tag{8.16}$$

Besides, it is easy to see that

$$|V_1| \leqslant \sum_{k=1}^{n_1} f_k + \sum_{j=1}^{n_2} f'_j, \tag{8.17}$$

because at least one parameter is fixed in each vertex from a deleted clique. Using these relations, similarly to the proof of Lemma 8.9, taking $\kappa = n_1 + n_2 < q$, we can write:

$$\begin{aligned}
\rho^{|V|}\bigl\|\mathrm{Prod}\bigl(X(G)\bigr)\bigr\|_p &\leqslant n^K \cdot \prod_{k=1}^{n_1} \rho^{|V(\widetilde{G}_k)|}\bigl\|\mathrm{Prod}\bigl(X(\widetilde{G}_k)\bigr)\bigr\|_{\kappa p} \cdot \prod_{j=1}^{n_2} \rho\|f_{\vec{r}_{u_j}}\|_{\kappa p} \\
&\lesssim n^K \cdot \prod_{k=1}^{n_1} \bigl[Cp^{\frac{d-1}{2}} q^d n^{-\frac{1}{16d}}\bigr]^{|V(\widetilde{G}_k)|} n^{-\frac{f_k}{2}} \cdot \prod_{j=1}^{n_2} \bigl[p^{\frac{d-1}{2}} q^d n^{-\frac{f'_j}{2}}\bigr] \\
&\lesssim \bigl[Cp^{\frac{d-1}{2}} q^d\bigr]^{|V|} \cdot n^{K - \frac{1}{2}(\sum f_k + \sum f'_j) - \frac{1}{16d}\sum |V(\widetilde{G}_k)|} \\
&\lesssim \bigl[Cp^{\frac{d-1}{2}} q^d\bigr]^{|V|} n^{-\frac{1}{6}|V_1| - \frac{1}{16d}|V_2|} \lesssim \bigl[Cp^{\frac{d-1}{2}} q^d n^{-\frac{1}{16d}}\bigr]^{|V|}.
\end{aligned}$$

9. The lower bound on the Discrepancy Function

We give the proof of Theorem 2.4, which is essentially a corollary to the proof of our main theorem, Theorem 1.5. As such, we will give a somewhat abbreviated proof. Indeed, the analogy between the lower bound on Discrepancy Functions and the Small Ball Inequality is well known to experts.

The proof is by duality. Fix N, and take $2N \leqslant 2^n < 4N$. It is a familiar fact [2] that for each $|\vec{r}| = n$ we can construct an r function $f_{\vec{r}}$ such that

$$\langle D_N, f_{\vec{r}}\rangle > c > 0, \tag{9.1}$$

where c depends only on dimension. We use these functions in the construction of the test function, following Section 4, with this one change. Before, see (4.3), we took $I_1, \dots, I_q$ to be a partition of $\{1, 2, \dots, n\}$ into q disjoint intervals of equal length. Instead, we take

$$I_t := \{j \in \mathbb{N}\colon |j - tn/q| < q/4\}. \tag{9.2}$$

This is the only change we make in the construction of Ψ^{sd}. It follows that $\|\Psi^{\mathrm{sd}}\|_1 \lesssim 1$.

Recall that $\Psi^{\mathrm{sd}} = \sum_{k=1}^{q} \Psi_k^{\mathrm{sd}}$, see (4.5). By construction, we have

$$\begin{aligned}\langle D_N, \Psi_1^{\mathrm{sd}}\rangle &= \sum_{t=1}^{q} \tilde{\rho} \sum_{\vec{r}\in\mathbb{A}_t} \langle D_N, f_{\vec{r}}\rangle \\ &\gtrsim q^b n^{(d-1)/2} \simeq n^{\epsilon/4+(d-1)/2}.\end{aligned}$$

This is a 'gain over the average case estimate' as one can see by comparison to Theorem 2.2. It remains to see that the higher order terms Ψ_k^{sd} contribute smaller terms than the one above.

By construction, Ψ_k^{sd} is itself a sum of r functions $f_{\vec{s}}$ with $|\vec{s}| > n$. Indeed, it follows from the separation in (9.2) that we necessarily have

$$n + k\frac{n}{2q} \leqslant |\vec{s}| \leqslant nd. \tag{9.3}$$

Second, it is a well-known fact that $|\langle D_N, f_{\vec{s}}\rangle| < N2^{-|\vec{s}|}$. Third, we fix $\vec{s}$ as above, and set $\mathrm{Count}(\vec{s})$ to be the number of distinct ways can we select $\vec{r}_1, \dots, \vec{r}_k$, all of length n, so that the product $f_{\vec{r}_1} \cdots f_{\vec{r}_k}$ is an r function of parameter $\vec{s}$. A very crude bound here is sufficient,

$$\mathrm{Count}(\vec{s}) \leqslant |\vec{s}|^{(d-1)k}.$$

Thus, we can estimate

$$\begin{aligned}\langle D_N, \Psi_k^{\mathrm{sd}}\rangle &\leqslant \sum_{j\geqslant n+k\frac{n}{2q}} \Bigl(\sum_{\vec{s}\colon |\vec{s}|=j} \mathrm{Count}(\vec{s})\bigl|\langle D_N, f_{\vec{s}}\rangle\bigr|\Bigr) \\ &\lesssim n^{d(k+3)} 2^{-kn/2q}.\end{aligned}$$

As $q = n^{\epsilon}$, this is clearly summable in $k \geqslant 1$ to at most a constant. This completes the proof.

10. The proof of the smooth Small Ball Inequality

We prove Theorem 2.7. There is no loss of generality in assuming that $|\alpha(R)| \leqslant 1$ for all R of volume at least 2^{-n}, since both sides of (2.8) are homogeneous and sums have finitely many terms. With φ as in the theorem, set

$$\varphi_{\vec{r}} = \sum_{R\colon |R_j|=2^{-r_j}} \alpha(R)\varphi_R.$$

And let $\Phi = \sum_{|\vec{r}|=n} \varphi_{\vec{r}}$. Define the r functions as in (3.7). It is the assumption that $c_\varphi = \langle \varphi, h_{[-1/2,1/2]} \rangle \neq 0$, and in fact we will assume that this inner product is positive. Thus,

$$\langle \varphi_{\vec{r}}, f_{\vec{r}} \rangle = c_\varphi 2^{-n} \sum_{R:\ |R_j|=2^{r_j}} \bigl|\alpha(R)\bigr|. \tag{10.1}$$

As $\varphi \in C[-1/2, 1/2]$, we have

$$\bigl|\langle \varphi, h_J \rangle\bigr| \leqslant C_\varphi |J| \tag{10.2}$$

for all dyadic intervals J.

It is important to note that

$$\bigl|\langle \varphi_{\vec{r}}, f_{\vec{s}} \rangle\bigr| \lesssim \begin{cases} 0 & \exists j\colon\ s_j < r_j, \\ C_\varphi 2^{-|\vec{r}-\vec{s}|} & \text{otherwise.} \end{cases} \tag{10.3}$$

The first line follows from the fact that φ is supported on $[-1/2, 1/2]$, so that if e.g. $s_1 < r_1$, the fact that φ has mean zero proves this estimate. The second estimate follows from (10.2) and the assumption that the coefficients $\alpha(R)$ are at most one in absolute value.

Let us take the intervals I_t in (9.2), and let us assume that

$$\sum_{|R|=2^n} \bigl|\alpha(R)\bigr| \leqslant 4 \sum_{t=1}^{q} \sum_{\vec{r}\in\mathbb{A}_t} \sum_{R:\ |R_j|=2^{-r_j}} \bigl|\alpha(R)\bigr|. \tag{10.4}$$

If this inequality fails, it is an easy matter to redefine the I_t so that the inequality above is true, and adjacent intervals I_t, I_{t+1} are separated by n/q.

We then follow Section 4 as before to define our test function Ψ^{sd}. It follows that $\|\Psi^{\mathrm{sd}}\|_1 \lesssim 1$. Using (10.4), (10.1) and (10.3), we have

$$\begin{aligned} \bigl\langle \Phi, \Psi_1^{\mathrm{test}} \bigr\rangle &\geqslant c 2^{-n} \tilde{\rho} \sum_{t=1}^{q} \sum_{\vec{r}\in\mathbb{A}_t} \sum_{R:\ |R_j|=2^{r_j}} \bigl|\alpha(R)\bigr| \\ &\gtrsim 2^{-n} n^{-(d-1)/2+\epsilon/4} \sum_{|R|=2^{-n}} \bigl|\alpha(R)\bigr|. \end{aligned}$$

This is the main term.

It remains to see that the inner products $|\langle \Phi, \Psi_k^{\mathrm{sd}} \rangle|$ are small $k \geqslant 1$. The details of this calculation are very similar to the corresponding calculations in the previous section, hence they are omitted.

References

[1] József Beck, A two-dimensional van Aardenne–Ehrenfest theorem in irregularities of distribution, Compos. Math. 72 (3) (1989) 269–339, MR1032337 (91f:11054).

[2] József Beck, William W.L. Chen, Irregularities of Distribution, Cambridge Tracts in Math., vol. 89, Cambridge University Press, Cambridge, 1987, MR903025 (88m:11061).

[3] Dmitriy Bilyk, Michael T. Lacey, On the Small Ball Inequality in three dimensions, Duke Math. J., (2006), in press arXiv: math.CA/0609815.

[4] Thomas Dunker, Estimates for the small ball probabilities of the fractional Brownian sheet, J. Theoret. Probab. 13 (2) (2000) 357–382, MR 1777539 (2001g:60085).

[5] Thomas Dunker, Thomas Kühn, Mikhail Lifshits, Werner Linde, Metric entropy of the integration operator and small ball probabilities for the Brownian sheet, C. R. Acad. Sci. Paris Sér. I Math. 326 (3) (1998) 347–352 (in English, with English and French summaries), MR2000b:60195.

[6] R. Fefferman, J. Pipher, Multiparameter operators and sharp weighted inequalities, Amer. J. Math. 119 (2) (1997) 337–369, MR1439553 (98b:42027).

[7] G. Halász, On Roth's method in the theory of irregularities of point distributions, in: Recent Progress in Analytic Number Theory, vol. 2, Durham, 1979, Academic Press, London, 1981, pp. 79–94, MR 637361 (83e:10072).

[8] James Kuelbs, Wenbo V. Li, Metric entropy and the small ball problem for Gaussian measures, J. Funct. Anal. 116 (1) (1993) 133–157, MR 94j:60078.

[9] Thomas Kühn, Werner Linde, Optimal series representation of fractional Brownian sheets, Bernoulli 8 (5) (2002) 669–696, MR 2003m:60131.

[10] W.V. Li, Q.-M. Shao, Gaussian processes: Inequalities, small ball probabilities and applications, in: Stochastic Processes: Theory and Methods, in: Handbook of Statistics, vol. 19, North-Holland, Amsterdam, 2001, pp. 533–597, MR 1861734.

[11] K.F. Roth, On irregularities of distribution, Mathematika 1 (1954) 73–79, MR 0066435 (,575c).

[12] Wolfgang M. Schmidt, Irregularities of distribution, VII, Acta Arith. 21 (1972) 45–50, 0319933 (47#8474).

[13] Wolfgang M. Schmidt, Irregularities of distribution. X, in: Number Theory and Algebra, Academic Press, New York, 1977, pp. 311–329, MR 0491574 (58#10803).

[14] Michel Talagrand, The small ball problem for the Brownian sheet, Ann. Probab. 22 (3) (1994) 1331–1354, MR 95k:60049.

[15] V.N. Temlyakov, Approximation of functions with bounded mixed derivative, Proc. Steklov Inst. Math. 1 (178) (1989), vi+121, MR 1005898 (90e:00007).

[16] V.N. Temlyakov, An inequality for trigonometric polynomials and its application for estimating the entropy numbers, J. Complexity 11 (2) (1995) 293–307, MR 96c:41052.

[17] Gang Wang, Sharp square-function inequalities for conditionally symmetric martingales, Trans. Amer. Math. Soc. 328 (1) (1991) 393–419, MR 1018577 (92c:60067).